



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 109507

TO: Karen A Lacourciere
Location: CM1/11D0911E12
Art Unit: 1635
Wednesday, December 03, 2003

Case Serial Number: 08765244

From: Barb O'Bryen
Location: Biotech-Chem Library
CM1-6A05
Phone: 308-4291 *BOB*

barbara.obryen@uspto.gov

Search Notes

RUSH

STIC-Biotech/ChemLib

109507

From: Fredman, Jeffrey
Sent: Tuesday, December 02, 2003 2:40 PM
To: STIC-Biotech/ChemLib
Cc: Lacourciere, Karen
Subject: FW: Rush sequence search request 08/765,244

PLEASE RUSH.

I Approve.

Jeff Fredman

CRFE

-----Original Message-----

From: Lacourciere, Karen
Sent: Tuesday, December 02, 2003 2:08 PM
To: Fredman, Jeffrey
Subject: Rush sequence search request 08/765,244

Hi Jeff-

Could you approve a RUSH for this search? It is a short amino acid sequence. The case is a 2 month amended due this biweek and I did not realize it had claims to a particular sequence.

Thanks-

Karen

Please search SEQ ID NO:22 in the commercial databases and pre-grant publications.
Thank-you!

Karen A. Lacourciere Ph.D.

CM1 11D09 GAU 1635

(703) 308-7523

mailbox 11E12

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	217	100.0		43	17	AAR90584	Rat ornithine tran
2	159	73.3		32	23	ABG30857	Rat ornithine tran
3	109	50.2		258	21	AA118445	Protein encoded by
4	108	49.8		32	23	ABG30856	Human ornithine tr
5	105	48.4		32	22	AAG54224	OTC peptide fragme
6	92	42.4		32	15	AAR48260	Mitochondrial matr
7	92	42.4		32	20	AAW96358	Mitochondrial matr
8	92	42.4		32	23	ABG32953	Localisation sequ
9	92	42.4		32	24	ABP56588	Mitochondrial matr

PT and molecular therapy of genetic diseases.

XX Disclosure; Column 11; 24pp; German.

XX Two modified oligonucleotides (introducing PstI and XhoI sites) were used to amplify a region of the human mitochondrial (mt) genome containing the light strand promoter, mtDNA ori of the heavy strand, CSB's ("conserved sequence blocks") and a regulation site for DNA replication. Behind this fragment (5' direction) a synthetic multiple cloning site was introduced, generating a product with overhang ends. The synthetic region also introduced a bidirectional mt transcription termination sequence. The amplification product, synthetic fragment and pBluescript were ligated and recombinant plasmid 1 (AAT12315) was produced. Human mt 16S rRNA (differing from the native RNA only in having a modified nucleotide) was isolated by PCR from chloramphenicol resistant HeLa cells and inserted into plasmid 1 to form plasmid 2 (AAT12316). The cloned insert was isolated as a BsaI fragment and cyclised using hairpin loop oligonucleotides, one of which carried the required signal peptide (the present sequence). The cyclised product was purified by treatment with exonuclease III. In a modification, the signal peptide was attached after cyclisation. The new plasmids were able to impart chloramphenicol resistance to otherwise sensitive B lymphocytes and fibroblasts. Similar plasmids without a signal peptide could not do this. The plasmids can be introduced into eukaryotic cells, esp. for site-directed mutagenesis or molecular therapy of genetic diseases, targeting nucleic acid in cells or their organelles via the protein transport route. (Updated on 25-MAR-2003 to correct PR field.)

XX Sequence 43 AA;

SQ Query Match 100.0%; Score 217; DB 17; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.1e-26;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSNRLILNKALRAKHAHTSMVRNFRYKPVQSQVQLKPRDLC 43
DB 1 MLSNRLILNKALRAKHAHTSMVRNFRYKPVQSQVQLKPRDLC 43

RESULT 2
ABG30857
ID ABG30857 standard; Peptide; 32 AA.
AC ABG30857;
DT 07-OCT-2002 (first entry)
XX Rat ornithine transcarbamylase signal peptide.
DE Rat; signal peptide; ornithine transcarbamylase; MOT;
KW recombinant vector; fusion protein; extranuclear gene.
XX Rattus sp.
OS JP2002176988-A.
PN 25-JUN-2002.
XX 14-DEC-2000; 2000JP-0380975.
PF 14-DEC-2000; 2000JP-0380975.
XX (TANAKA) TANAKA M.
PA (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUJO.
PA (OYOS-) ZH OYO SEIKAGAKU KENKYUSHO.
XX WPI; 2002-569946/61.
DR N-PSDB; ABX88419.
XX A recombinant vector for expressing a fused protein, useful for decomposing an extranuclear gene of a nonhuman organism -
PT Disclosure; Page 3; 15pp; Japanese.

XX The invention relates to a recombinant vector for expressing a fused protein containing a fused gene in which a base sequence defining a transfer signal peptide to small cellular organs having an extranuclear gene is combined with a base sequence defining the amino acid sequence of a restriction enzyme recognising a defined base sequence. The vector is used for decomposing an extranuclear gene of a nonhuman organism. The present sequence represents the signal peptide of rat mitochondrial ornithine transcarbamylase (MOT) which may be used in the vector of the invention.

XX Sequence 32 AA;

SQ Query Match 73.3%; Score 159; DB 23; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.4e-17;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLSNRLILNKALRAKHAHTSMVRNFRYKPVQ 32
DB 1 MLSNRLILNKALRAKHAHTSMVRNFRYKPVQ 32

RESULT 3
AAB18445
ID AAB18445 standard; Protein; 258 AA.
XX AAB18445;
AC AAB18445;
XX 15-JAN-2001 (first entry)
DT Protein encoded by plasmid pUOATP2.
DE Mitochondrial genome; mitochondrial encephalomyopathy; lactic acidosis; Leber hereditary optic neuropathy; myoclonic epilepsy;
KW neurogenic muscular weakness; ataxia; retinitis pigmentosa;
KW Kearns-Sayre syndrome; Leigh syndrome; Pearson Marrow pancreas syndrome; aminoglycoside-associated deafness; diabetes; deafness; leukodystrophy;
KW hypotonia; autism; sudden infant death syndrome; hypoglycemia; leukaemia; thrombocytopenia; migraine; hearing loss; stroke;
KW refractory infantile reflux; carnitine deficiency; multiple sclerosis;
KW blindness; optic atrophy; renal tubular acidosis; cardiomyopathy;
KW chronic pancreatitis; ATPase 6.
XX Synthetic.
OS WO200053773-A2.
PN 14-SEP-2000.
XX 08-MAR-2000; 2000WO-US06037.
PF 08-MAR-1999; 99US-0123336.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Zullo SJ, Eisenstadt JM;
PI WPI; 2000-565602/52.
XX N-PSDB; AAA75084.
DR Functionally complementing one or more defects, mutations, or deletions in a mitochondrial genome of a cell, useful for treating diabetes with deafness, comprises introducing mitochondrial DNA into the nuclear genome -
PT Example 1; Fig 3; 49pp; English.
PS The specification describes a method for functionally complementing one or more defects, mutations, or deletions in a mitochondrial genome of a cell having a nuclear genome. The method is used for treating a disease or disorder that arises from deletion of the protein-encoding genes of the mitochondrial genome. Alternatively, the method is used for treating a disease or disorder that arises from one or more defects, deletions or

CC mutations in mitochondrial genes encoding ribosomes or tRNA for
 CC transcription and translation in the mitochondria. The diseases or
 CC disorders that can be ameliorated are mitochondrial encephalomyopathy with
 CC lactic acidosis and stroke-like episodes, Leber hereditary optic
 CC neuropathy, myoclonic epilepsy with ragged-red fibres, neurogenic
 CC muscular weakness, ataxia, retinitis pigmentosa, Kearns-Sayre syndrome,
 CC Leigh syndrome, Pearson Marrow pancreas syndrome, aminoglycoside-
 CC associated deafness, diabetes with deafness, leukodystrophy with
 CC hypotonia, autism with seizures, sudden infant death syndrome with
 CC hypoglycemia, leukaemia with maternally inherited thrombocytopenia,
 CC migraines (associated with hearing loss, strokes, or diabetes), early
 CC hearing loss, refractory infantile reflux with carnitine deficiency,
 CC multiple sclerosis with seizures, blindness with optic atrophy and
 CC dystonia, renal tubular acidosis with elevated lactic acid and hypotonia,
 CC nonvalvular hypertrophic cardiomyopathy before age 50 and chronic
 CC pancreatitis with stroke-like episodes. The present sequence is encoded
 CC by plasmid pHOATP2. Plasmid pHOATP2 comprises a mutant oligomycin-
 CC resistant ATPase 6 mitochondrial gene derived from Chinese hamster ovary
 CC (CHO) cells linked to ornithine transcarbamylase DNA sequence. It is
 CC used for targeting the protein to the mitochondria.

CC Sequence 258 AA;
 Query Match 50.2%; Score 109; DB 21; Length 258;
 CC Best Local Similarity 62.2%; Pred. No. 2e-08;
 CC Matches 23; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALKAKHTSMVRNFRYKPKVQSQVL 37
 DB 1 MLFNRLILNNAAFRNGHFMVRNFRGCGQLQMNENL 37

RESULT 4
 ABG30856
 ID ABG30856 standard; Peptide; 32 AA.

XX AC ABG30856;

DT 07-OCT-2002 (first entry)

XX Human ornithine transcarbamylase signal peptide.

XX Human; signal peptide; ornithine transcarbamylase; MOT;
 KW recombinant vector; fusion protein; extranuclear gene.

XX Homo sapiens.

XX JP2002176988-A.

XX (GIFU-) ZH GIFUKEN KOKUSAI BIO KENYUO.

XX (OYOS-) ZH OYO SEIKAGAKU KENKYUSHO.

XX WPI; 2002-569346/61.

XX N-PSDB; ABK88418.

XX A recombinant vector for expressing a fused protein, useful for
 PT decomposing an extranuclear gene of a nonhuman organism -
 PS Disclosure; Page 3; 15pp; Japanese.

CC The invention relates to a recombinant vector for expressing a fused
 CC protein containing a fused gene in which a base sequence defining a
 CC transfer signal peptide to small cellular organs having an extranuclear
 CC gene is combined with a base sequence defining the amino acid sequence of
 CC a restriction enzyme recognising a defined base sequence. The vector is
 CC used for decomposing an extranuclear gene of a nonhuman organism.

CC The present sequence represents the signal peptide of human mitochondrial
 CC ornithine transcarbamylase (MOT) which may be used in the vector of the
 CC invention.

XX Sequence 32 AA;

Query Match 49.8%; Score 108; DB 23; Length 32;

CC Best Local Similarity 68.8%; Pred. No. 2.2e-09;

XX Matches 22; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALKAKHTSMVRNFRYKPKVQ 32

DB 1 MLFNRLILNNAAFRNGHFMVRNFRGCGQLQ 32

RESULT 5

AAG64224

ID AAG64224 standard; Peptide; 32 AA.

XX AC AAG64224;

XX DT 19-SEP-2001 (first entry)

XX OTC peptide fragment.

XX KW Heat shock protein interacting protein; HSP47; OTC.

XX OS Unidentified.

XX PN JP2001145493-A.

XX PD 29-MAY-2001.

XX PF 19-NOV-1999; 99JP-0330631.

XX PR 19-NOV-1999; 99JP-0330631.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX WPI; 2001-395263/42.

XX Using the two-hybrid screening method to prepare proteins which
 PT interact with the heat shock protein HSP47 -

XX Example 7; Page 12; 26pp; Japanese.

CC The present invention relates to a method for preparing a protein which
 CC interacts with the heat shock protein HSP47. The method involves the
 CC two-hybrid screening method using the HSP47 gene and a mammalian cDNA
 CC library. The HSP47 interacting proteins are useful for the diagnosis and
 CC treatment of diseases caused by an increase or decrease in activity of
 CC HSP47. The present sequence was used in an example from the present
 CC invention.

XX Sequence 32 AA;

Query Match 48.4%; Score 105; DB 22; Length 32;

CC Best Local Similarity 68.8%; Pred. No. 6.4e-09;

XX Matches 22; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALKAKHTSMVRNFRYKPKVQ 32

DB 1 MLFNRLILNNAAFRNGHFMVRNFRGCGQLQ 32

RESULT 6

AAR48260

ID AAR48260 standard; peptide; 32 AA.

XX AC AAR48260;

XX DT 25-MAR-2003 (updated)

XX DT 29-JUL-1994 (first entry)

XX Mitochondrial matrix retention signal.
 XX Single chain antibody; sFv; heavy chain; light chain;
 KW variable domain; hydrophilic linker; antibodies; targeting;
 KW subcellular localisation signal; mitochondrial matrix;
 XX retention signal.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 7 /note= "not defined"
 FT Misc-difference 8 /note= "not defined"
 FT Misc-difference 32 /note= "not defined"
 FT Misc-difference 32 /note= "not defined"
 XX WO9402610-A1.
 FN XX
 PD 03-FEB-1994.
 PD 16-JUL-1993; 93WO-US06735.
 PF 17-JUL-1992; 92US-0916939.
 PR 17-MAR-1993; 93US-0045274.
 PR XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Haseltine WA, Marasco WA;
 XX WPI; 1994-048868/06.
 DR Intracellular binding of antigens - by using antibody targeting
 PT with vector system, for e.g. tumour suppression
 PT
 XX Disclosure; Page 103; 155pp; English.
 PS New vector systems comprise a sequence adapted for intracellular
 CC delivery and expression congt. a promoter operably linked to an
 CC antibody gene encoding an antibody which binds to a specific target
 CC antigen. The antibody is esp. a single chain antibody in which the
 CC heavy and light chain variable regions are joined via a hydrophilic
 CC linker peptide. Localisation sequences are pref. included in the
 CC constructs. The sequence AA#48260 is a mitochondrial matrix retention
 CC signal.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX Sequence 32 AA;
 SQ Query Match 42.4%; Score 92; DB 15; Length 32;
 Best Local Similarity 61.3%; Pred. No. 6.8e-07;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MLSNRLILNKALRKAHTSMVRNFRYKPV 31
 |||||
 Db 1 MLEFNLRXXLNNAAFRGHNFVRNFRGQPL 31
 |||||
 RESULT 7
 AAW96358
 ID AAW96358 standard; Peptide; 32 AA.
 XX AAW96358;
 AC AAW96358;
 XX 19-JUL-1999 (first entry)
 DT Mitochondrial matrix localisation signal peptide.
 XX Antibody; immune response; modulation; MHC; IRM; receptor;
 KW intrabody; major histocompatibility complex; graft rejection;
 KW immunomodulatory response molecule; regulation; transplantation;
 KW retention signal; localisation signal; golgi apparatus; ER;

KW endoplasmic reticulum.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 7 /note= "Any amino acid"
 FT Misc-difference 8 /note= "Any amino acid"
 FT Misc-difference 32 /note= "Any amino acid"
 FT Misc-difference 32 /note= "Any amino acid"
 XX WO9914353-A2.
 PN 25-MAR-1999.
 PD 18-SEP-1998; 98WO-US19563.
 PF 19-SEP-1997; 97US-0059339.
 PR (DAND) DANA FARBER CANCER INST INC.
 XX Marasco W, Mhashikar A;
 PI WPI; 1999-229546/19.
 DR Altering the regulation of the immune system
 XX Disclosure; Page 28; 56pp; English.
 PS Intracellular binding to a desired target by an intracellularly
 CC expressed antibody (i.e. an intrabody) can be used to knock out
 CC multiple locuses of immunomodulatory receptor molecules (IRMs),
 CC so that the expression of multiple major histocompatibility (MHC)
 CC molecules is blocked. This selective targeting of IRMs, their
 CC pathways or components, can be used to selectively regulate the
 CC immune system by controlling expression of these molecules and
 CC preventing an undesired immune response in a cell. Any component
 CC of the MHC pathway or the MHC assembly line or antigen presentation
 CC can be targeted. Intrabodies can be used to knock out the immune
 CC response in a particular tissue or portion of the body to prepare
 CC it for cell or tissue transplantation. Alternatively, an organ for
 CC transplantation can be perfused with the intrabody ex vivo. The
 CC intrabodies can comprise whole antibodies, heavy chains, Fab'
 CC fragments, single-chain antibodies and diabodies. The intrabodies
 CC also comprise an intracellular localisation signal to facilitate
 CC interception of expressed proteins. For example, if the target was
 CC a cell surface receptor, the antibody would comprise a leader
 CC sequence and an endoplasmic reticulum (ER) or Golgi apparatus
 CC retention signal. This peptide is a localisation sequence for the
 CC mitochondrial matrix. For other localisation sequences see
 CC AAW96345-W96377.
 XX Sequence 32 AA;
 SQ Query Match 42.4%; Score 92; DB 20; Length 32;
 Best Local Similarity 61.3%; Pred. No. 6.8e-07;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
 QY 1 MLSNRLILNKALRKAHTSMVRNFRYKPV 31
 |||||
 Db 1 MLEFNLRXXLNNAAFRGHNFVRNFRGQPL 31
 |||||
 RESULT 8
 ABG92993
 ID ABG92993 standard; Peptide; 32 AA.
 XX ABG92993;
 AC ABG92993;
 DT 20-NOV-2002 (first entry)
 XX Localisation sequence to direct antibodies to the mitochondria.
 DE

XX Regulator; transcription; cell death; phenotype; molecular scaffold;
 KW gene therapy; cancer; cardiovascular disease; arrhythmia; heart failure;
 KW ischaemia; obesity; neurodegenerative disease; Alzheimer's disease;
 KW bone pathology; dermatologic disease; psoriasis; infection; AIDS;
 KW acquired immunodeficiency syndrome; cosmetic; wound healing;
 KW antibiotic transport; drug toxicity; drug resistance; immunobiology;
 KW inflammation; allergic response; human immunodeficiency virus.
 XX Unidentified.
 OS
 XX
 XX WO200262822-A2.
 PN
 XX
 XX 15-AUG-2002.
 PD
 XX
 XX 04-FEB-2002; 2002WO-USQ2814.
 PF
 XX
 XX 02-FEB-2001; 2001US-265589P.
 PR
 XX 05-FEB-2001; 2001US-265880P.
 PR
 XX 27-FEB-2001; 2001US-271423P.
 PR
 XX 23-JAN-2001; 2001US-263226P.
 PR
 XX 28-MAR-2000; 2000US-192586P.
 PR
 XX 22-SEP-1997; 97US-935377P.
 PR
 XX
 XX (UYRP) UNIV ROCHESTER.
 PA
 XX
 XX Zauderer M, Smith ES;
 PI
 XX
 XX WPI; 2002-643398/69.
 DR
 XX
 XX Identifying regulator polypeptides which influence target
 PT transcriptional regulatory regions, useful for treating cancer.
 PT comprises introducing host cells expressing the polypeptide into a
 PT library of polynucleotides -
 XX
 XX Disclosure; Page 37; 224pp; English.
 PS
 XX
 XX The invention discloses a method for identifying polynucleotides encoding
 CC a regulator polypeptide, whose expression induces activation of a target
 CC transcriptional regulatory region in a host cell. The method comprises
 CC providing a population of eukaryotic host cells capable of expressing the
 CC polypeptide, introducing into the host cell a library of polynucleotides
 CC encoding the polypeptides, permitting expression of the polypeptides and
 CC then recovering them from the host cells. The target transcriptional
 CC regulatory region is operably associated with a polynucleotide encoding a
 CC gene product the expression of which results in host cell death or cause
 CC the host cells to exhibit a pre-determined modified phenotype and where
 CC the gene product is expressed upon activation of target transcriptional
 CC regulatory region. Each candidate regulator polypeptide comprises a
 CC candidate peptide and a molecular scaffold fused to the peptide so that
 CC the peptide is displayed on the surface of the candidate regulator
 CC polypeptide. The methods are useful in selecting and/or screening
 CC regulator molecules, such as polypeptides, which directly or indirectly
 CC induce or suppress the transcriptional activation of a target
 CC transcriptional regulatory region in a eukaryotic host cell. These
 CC regulator molecules may be used (e.g. in gene therapy) for preventing or
 CC treating cancers (e.g. breast or ovarian cancer), cardiovascular diseases
 CC (e.g. arrhythmia, heart failure, ischaemia), obesity, neurodegenerative
 CC diseases (e.g. Alzheimer's disease), bone pathologies, dermatologic
 CC diseases (e.g. psoriasis), infections (e.g. viral, bacterial), acquired
 CC immunodeficiency syndrome (AIDS), in cosmetic applications and in wound
 CC healing. The method is also useful in screening regulator molecules that
 CC block antibiotic transport mechanisms, in drug toxicities and drug
 CC resistance applications and in improving the performance of existing or
 CC developmental drugs. It may also be used in immunobiology, inflammation,
 CC allergic response and in biotechnology applications. The sequences
 CC presented in ABG92946-ABG93029 are examples of regulator polypeptides.
 XX
 XX Sequence 32 AA;
 SQ
 Query Match 42.4%; Score 92; DB 23; Length 32;
 Best Local Similarity 61.3%; Pred. No. 6.8e-07;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRIILINKAALKAKHTSMVRNPRYKPV 31
 |||||
 DB 1 MLENLXXXLNNAAPRHGHNFVNRFCQQL 31
 |||||

RESULT 9

ABP56588
 ID ABP56588 standard; Peptide; 32 AA.

XX AC ABP56588;

XX DT 24-MAR-2003 (first entry)

XX DE Mitochondrial matrix targeting peptide SEQ ID NO:54.

XX Identification; intrabody; eukaryotic cell; immunoglobulin; selection;
 KW cardiovascular; diminished arrhythmia potential; cardiomyocyte; stroke;
 KW enhanced contractile property; heart failure; arrhythmia; embolic;
 KW sarcolemmal calcium cycling; artery; arteriole; angina; atherosclerosis;
 KW LDL metabolism; HDL metabolism; skin biology; keloid formation.

XX OS Unidentified.

XX PH Key Location/Qualifiers

FT Misc-difference 7 /note= "any amino acid"

FT Misc-difference 8 /note= "any amino acid"

FT Misc-difference 32 /note= "any amino acid"

XX PN WO200286096-A2.

XX 31-OCT-2002.

XX 23-JAN-2002; 2002WO-USQ1677.

XX 23-JAN-2001; 2001US-263225P.

XX 24-JAN-2001; 2001US-263200P.

XX 27-FEB-2001; 2001US-271422P.

XX 15-JUN-2001; 2001US-298095P.

XX (UYRP) UNIV ROCHESTER MEDICAL CENT.

XX Zauderer M, Wei C, Smith ES;

XX WPI; 2003-103408/09.

XX Selecting polynucleotides encoding an intracellular immunoglobulin
 PT which induces a modified phenotype in a eukaryotic host cell, by
 PT introducing library of polynucleotides encoding immunoglobulin subunit
 PT polypeptides -

XX Disclosure; Page 44; 257pp; English.

PS The present invention describes a method for selecting polynucleotides
 CC (PNS) encoding an intracellular immunoglobulin molecule or its fragment
 CC whose expression induces a modified phenotype in a eukaryotic host cell
 CC (I). The method comprises introducing into (i) a first and second library
 CC of PNS encoding, through operable association with a transcriptional
 CC control region, first and second intracellular immunoglobulin subunit
 CC polypeptides, respectively. The method is useful for selecting
 CC polynucleotides which encode an intracellular immunoglobulin molecule, or
 CC fragment. The method is useful e.g. for identifying polynucleotides which
 CC singly or collectively encode intracellular immunoglobulin molecules, or
 CC which sensitize host cells to killing by an agent. The method may also be
 CC used in cardiovascular applications; for screening for diminished
 CC arrhythmia potential in cardiomyocytes and for enhanced contractile
 CC properties of cardiomyocytes and diminish heart failure potential; for
 CC identifying intracellular immunoglobulin molecules that will regulate
 CC intracellular and sarcolemmal calcium cycling in cardiomyocytes to
 CC prevent arrhythmias or that will diminish embolic phenomena in arteries

ABB59259	
ID	ABB59259 standard; Protein; 755 AA.
XX	
AC	ABB59259;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 4569.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
XX	
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
PR	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE) PE CORP NY.
XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
DR	N-PSDB; ABL03362.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell interactions -
PT	
XX	
PS	Disclosure; SEQ ID NO 4569; 2lpp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling, cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
CC	
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the inventor at ftp.wipo.int/pub/published_pct_sequences .
XX	
SQ	Sequence 755 AA;

Query Match	24.7%	Score 53.5;	DB 22;	Length 755;
Best Local Similarity	42.5%	Pred. No. 33;		
Matches 17;	Conservative	5;	Mismatches 15;	Indels 3;
				Gaps 2;

Qy 2 LSNLRILLNKAAALRKAHTSMVRNFRYGVQSPVQSQVQLKPRD 41
||:||||| : : ||| | : : |||||
Db 11 LSDLRRELE--VLRKAHFEELDLFYGTG-OPEAEAKPRD 47

RESULT 15
 AAY04933
 ID AAY04933 standard; Protein; 165 AA.

AC AAY04933;

DT 06-JUL-1999 (first entry)

DE Mycobacterium species protein sequence 38B.

Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
hybridisation; detection; vaccine; immunisation; infection.
XX
OS Mycobacterium sp.

```

XX WO9909186-A2.
PN XX
XX XX
PD PD
XX 25-FEB-1999.
XX XX
XX 14-AUG-1998; 98WO-FR01813.
XX XX
XX 11-SEP-1997; 97FR-0011325.
PR PR
XX 14-AUG-1997; 97FR-0010404.
XX XX
XX (INSP ) INST PASTEUR.
PA PA
XX Gicquel B, Lim EM, Pellic V, Portnoi D, Goguet de la Salmoniere Y;
PI Guigueno A;
PI PI
XX WPI; 1999-181045/15.
DR DR
DR N-PSDB; AAX34186.
XX XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression
PT PT
XX Claim 32; Fig 38B; 309pp; French.
PS PS
XX Sequences AAY04742-Y05000 and AAY07201-Y07204 represent secreted
CC proteins from various Mycobacterium species microorganisms. The
CC encoding nucleotide sequences can be used as primers and probes for
CC methods for detecting and identifying mycobacteria, especially belonging
CC to the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.
XX XX
SQ Sequence 165 AA;

Query Match 24.0%; Score 52; DB 20; Length 165;
Best Local Similarity 37.1%; Pred. No. 8.9;
Matches 13; Conservative 8; Mismatches 12; Indels 2; Gaps

QY 6 RILNLKAALRKHAHTSMVRNFRYGKPVQSOLKPR 40
::|::|||::|::|::|::|::|
Db 119 RVILRCNHTKANKQSARTLRPL--RIALRP 151

Search completed: December 3, 2003, 14:32:33
Job time : 42 secs

```

Search completed: December 3, 2003, 14:32:33
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:31:47 : Search time 22 Seconds
(without alignments)
82.698 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MSLNLRILLNKAALRKXHTS.....NFRYKFPVQVQLKPRDLC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	42.4	32	2	US-08-373-190-29
2	92	42.4	32	2	US-08-438-190A-29
3	92	42.4	32	3	US-08-350-215-29
4	92	42.4	32	3	US-09-287-145A-29
5	92	42.4	32	4	US-09-556-111-29
6	54	24.9	2539	3	US-09-413-814-42
7	53	24.4	457	4	US-09-252-991A-32812
8	52	24.0	218	4	US-09-252-991A-18053
9	51	23.5	1169	1	US-08-542-921-2
10	51	23.5	1169	2	US-08-880-685-2
11	51	23.5	1169	2	US-08-880-684-2
12	49.5	22.8	208	4	US-09-107-332A-5399
13	49.5	22.8	523	4	US-09-252-991A-18693
14	49.5	22.8	1164	4	US-09-457-708-2
15	49.5	22.8	1164	4	US-09-950-046A-2
16	48	22.1	89	4	US-09-732-210-98
17	47	21.7	20	5	PCT-US95-07543-5
18	47	21.7	91	4	US-09-732-210-195
19	47	21.7	421	3	US-09-002-567B-1
20	47	21.7	421	3	US-09-002-567B-3
21	47	21.7	421	4	US-09-571-347-1
22	47	21.7	421	4	US-09-571-347-3
23	47	21.7	567	3	US-09-188-811-2
24	46.5	21.4	274	4	US-09-252-991A-29653
25	46.5	21.4	582	4	US-09-091-725-17
26	46.5	21.4	659	4	US-09-562-737-16
27	46.5	21.4	1711	3	US-08-369-822C-10

28 46.5 21.4 1711 3 US-08-582-776C-10 Sequence 10, Appl
29 46.5 21.4 1711 3 US-08-434-831B-10 Sequence 10, Appl
30 46 21.2 559 1 US-08-313-553-15 Sequence 15, Appl
31 46 21.2 559 3 US-08-767-993-15 Sequence 15, Appl
32 46 21.2 729 4 US-09-252-991A-32535 Sequence 32535, A
33 46 21.2 1139 4 US-09-914-259-68 Sequence 68, Appl
34 45.5 21.0 761 4 US-09-585-858-27 Sequence 27, Appl
35 45 20.7 236 4 US-09-252-991A-28901 Sequence 28901, A
36 45 20.7 275 4 US-09-328-352-4817 Sequence 14, Appl
37 45 20.7 535 4 US-09-312-762A-14 Sequence 14, Appl
38 45 20.7 609 4 US-09-107-532A-5110 Sequence 5110, Ap
39 45 20.7 700 4 US-08-671-757A-10 Sequence 10, Appl
40 45 20.7 716 4 US-09-219-983A-20 Sequence 20, Appl
41 45 20.7 730 1 US-07-846-181-5 Sequence 5, Appl
42 45 20.7 730 1 US-07-845-989-5 Sequence 5, Appl
43 45 20.7 795 4 US-09-252-991A-19085 Sequence 19085, A
44 45 20.7 984 4 US-09-328-352-6926 Sequence 6926, Ap
45 45 20.7 1124 4 US-09-252-991A-26810 Sequence 26810, A

ALIGNMENTS

RESULT 1
US-08-373-190-29
; Sequence 29, Application US/08373190
; Patent No. 5851829
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; APPLICANT: HASELTINE, WILLIAM
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF PROTEINS
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/373,190
; FILING DATE: 17-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06735
; FILING DATE: 16-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, DAVID S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 41956-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEFAX: STRE UR 2002
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-373-190-29

Query Match 42.4%; Score 92; DB 2; Length 32;

```
Best Local Similarity 61.3%; Pred. No. 9.2e-08;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKHAFTSMVNFYRGKPV 31
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHFMVNFRCGQPL 31

RESULT 2
US-08-438-190A-29
; Sequence 29, Application US/08438190A
; Patent No. 5965371
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSER: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/438,190A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-438-190A-29

Query Match 42.4%; Score 92; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 9.2e-08;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKHAFTSMVNFYRGKPV 31
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHFMVNFRCGQPL 31

RESULT 3
US-08-350-215-29
; Sequence 29, Application US/08350215
; Patent No. 6004940
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE A.
; TITLE OF INVENTION: INTRACELLULAR TARGETING OF ENDOGENOUS
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSER: CUSHMAN
; STREET: 130 WATER STREET

Query Match 42.4%; Score 92; DB 2; Length 32;
Best Local Similarity 61.3%; Pred. No. 9.2e-08;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKHAFTSMVNFYRGKPV 31
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHFMVNFRCGQPL 31

RESULT 4
US-09-287-145A-29
; Sequence 29, Application US/09287145A
; Patent No. 6072036
; GENERAL INFORMATION:
; APPLICANT: MARASCO, WAYNE
; TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSER: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/287,145A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/08/438,190
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: EISENSTEIN, RONALD I.
; REGISTRATION NUMBER: 30628
; REFERENCE/DOCKET NUMBER: 41956
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-350-215-29

Query Match 42.4%; Score 92; DB 3; Length 32;
Best Local Similarity 61.3%; Pred. No. 9.2e-08;
Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKHAFTSMVNFYRGKPV 31
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 MLFNLRXXLNNAAFRGHFMVNFRCGQPL 31
```

TELEFAX: (617) 523-6440
 TELEX: 200291 STRE UR
 INFORMATION FOR SEQ ID NO: 29:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 32 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 US-09-287-145A-29

Query Match 42.4%; Score 92; DB 3; Length 32;
 Best Local Similarity 61.3%; Pred. No. 9.2e-08;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAHTSMVRNFRYKGPV 31
 |||||
 DB 1 MLFNLRXXLNNAFRHHNFMVNRFCGQPL 31
 |||||

RESULT 5

US-09-556-111-29
 Sequence 29, Application US/09556111
 Patent No. 6329173

GENERAL INFORMATION:

APPLICANT: MARASCO, WAYNE
 HASELTINE, WILLIAM
 TITLE OF INVENTION: METHOD OF INTRACELLULAR BINDING OF
 PROTEINS

NUMBER OF SEQUENCES: 78

CORRESPONDENCE ADDRESS:

ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
 CUSHMAN

STREET: 130 WATER STREET

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: US

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/556,111

FILING DATE: 21-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/438,190

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: EISENSTEIN, RONALD I.

REGISTRATION NUMBER: 30628

REFERENCE/DOCKET NUMBER: 41956

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 523-3400

TELEFAX: (617) 523-6440

TELEX: 200291 STRE UR

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-09-556-111-29

Query Match 42.4%; Score 92; DB 4; Length 32;
 Best Local Similarity 61.3%; Pred. No. 9.2e-08;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLILNKALRKAHTSMVRNFRYKGPV 31
 |||||
 DB 1 MLFNLRXXLNNAFRHHNFMVNRFCGQPL 31
 |||||

RESULT 6

US-09-413-814-42
 Sequence 42, Application US/09413814
 Patent No. 6225064

GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
 APPLICANT: Bristol-Myers Squibb, Co.

APPLICANT: Beyer, Stefan

APPLICANT: Bloecker, Helmut

APPLICANT: Brandt, Petra

APPLICANT: Cino, Paul M

APPLICANT: Dougherty, Brian A

APPLICANT: Goldberg, Steven L

APPLICANT: Hofie, Gerhard

APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans

FILE REFERENCE: PCT/US 99/23535

TITLE OF INVENTION: heteropolyketide compounds

FILE REFERENCE: PCT/US 99/23535

CURRENT APPLICATION NUMBER: US/09/413,814

CURRENT FILING DATE: 1999-10-07

EARLIER APPLICATION NUMBER: DE 198 46 493.2

EARLIER FILING DATE: 1998-10-09

NUMBER OF SEQ ID NOS: 107

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 42

LENGTH: 2539

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-413-814-42

Query Match

24.9%; Score 54; DB 3; Length 2539;

Best Local Similarity 35.5%; Pred. No. 24;

Matches 11; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 12 AALRKAHTSMVRNFRYKGPVQSQVQLKPRDL 42
 |||||

DB 1969 ASLKAHSSPEPARHGRPALSSSEWVAPRV 1999
 |||||

RESULT 7

US-09-252-991A-32812

Sequence 32812, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 32812

LENGTH: 457

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

FEATURE:

NAME/KEY: UNSURE

LOCATION: (451)

OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-32812

Query Match

24.4%; Score 53; DB 4; Length 457;

Best Local Similarity 34.4%; Pred. No. 4.1;

Matches 11; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 7 ILNKALRKAHTSMVRNFRYKGPVQSQVQLK 38
 |||||

DB 176 VALDAVRRARLRVFRGLRPAQLYVETR 207
 |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:32:37 ; Search time 30 Seconds
(without alignments)
265.577 Million cell updates/sec

Title: US-08-765-244-22
Perfect score: 217
Sequence: 1 MLSNRIILNKALRKAHTS.....NFRYKPVQSQVQLKPRDLC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	100.0	43	8	US-08-765-244-22
2	197	90.8	41	8	Sequence 22, Appl
3	92	42.4	32	14	US-08-765-244-1
4	92	42.4	32	15	US-10-061-395-48
5	61	28.1	1286	11	US-10-052-942-54
6	61	28.1	1286	12	US-09-998-027-3
7	50.5	23.3	2212	12	US-10-165-099-3
8	50	23.0	414	15	US-10-353-856-8
9	50	23.0	564	12	US-10-156-761-7915
10	49.5	22.8	1164	10	US-10-032-585-7102
11	48.5	22.4	75	9	US-09-950-046A-2
12	48.5	22.4	285	9	US-09-864-761-46024
13	48.5	22.4	334	9	US-09-815-242-4869
14	48.5	22.4	684	10	US-09-815-242-10798
15	48	22.1	235	11	US-09-738-626-3622
					Sequence 3622, Ap
					Sequence 4156, Ap

16	48	22.1	618	12	US-10-061-201-7	Sequence 7, Appli
17	48	22.1	729	12	US-10-061-201-3	Sequence 3, Appli
18	48	22.1	1987	12	US-10-032-585-7518	Sequence 7518, Ap
19	47.5	21.9	826	12	US-09-756-247-41	Sequence 41, Appli
20	47	21.7	421	12	US-10-357-627-1	Sequence 1, Appli
21	47	21.7	421	12	US-10-357-627-3	Sequence 3, Appli
22	46.5	21.4	659	15	US-10-211-962-16	Sequence 16, Appli
23	46	21.2	1139	12	US-10-080-608A-68	Sequence 68, Appli
24	46	21.2	1139	12	US-10-370-685-157	Sequence 157, App
25	46	21.2	1551	9	US-09-864-761-35904	Sequence 35904, A
26	45.5	21.0	187	9	US-09-811-284-217	Sequence 217, App
27	45.5	21.0	513	15	US-10-177-293-98	Sequence 98, Appli
28	45.5	21.0	761	15	US-10-270-875-27	Sequence 27, Appli
29	45.5	21.0	761	15	US-10-270-878-27	Sequence 27, Appli
30	45.5	21.0	761	15	US-10-270-786-27	Sequence 27, Appli
31	45.5	21.0	761	15	US-10-370-710-27	Sequence 27, Appli
32	45.5	21.0	761	15	US-10-370-859-27	Sequence 27, Appli
33	45.5	21.0	761	16	US-10-270-846-27	Sequence 27, Appli
34	45.5	21.0	808	15	US-10-156-761-10993	Sequence 27, Appli
35	45.5	21.0	920	15	US-10-147-026-6	Sequence 6, Appli
36	45	20.7	511	10	US-09-881-752A-110	Sequence 110, App
37	45	20.7	511	11	US-09-988-067B-92	Sequence 92, Appli
38	45	20.7	535	10	US-09-312-762A-14	Sequence 14, Appli
39	45	20.7	700	12	US-09-015-078-10	Sequence 10, Appli
40	45	20.7	700	12	US-10-238-977A-10	Sequence 10, Appli
41	45	20.7	716	12	US-10-114-774-20	Sequence 20, Appli
42	45	20.7	730	15	US-10-099-352-42	Sequence 42, Appli
43	44.5	20.5	439	14	US-10-074-279-2	Sequence 2, Appli
44	44	20.3	103	11	US-09-969-730-125	Sequence 125, App
45	44	20.3	104	11	US-09-774-639-164	Sequence 164, App

ALIGNMENTS

RESULT 1
US-08-765-244-22
; Sequence 22, Application US/08765244
; Publication No. US20010008771A1
; GENERAL INFORMATION:
; APPLICANT: Seibel, Peter
; APPLICANT: Seibel, Andrea
; TITLE OF INVENTION: CHEMICAL PEPTIDE-NUCLEIC ACID
; TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR
; TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES
; TITLE OF INVENTION: AND CELLS
; FILE REFERENCE: 8484-0018-999
; CURRENT APPLICATION NUMBER: US/08/765,244
; CURRENT FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: PCT/DE95/00775
; PRIOR FILING DATE: 1995-06-11
; PRIOR APPLICATION NUMBER: DE P 44 21 079.5
; PRIOR FILING DATE: 1994-06-16
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Rattus rattus
US-08-765-244-22

Query Match 100.0%; Score 217; DB 8; Length 43;
Best Local Similarity 100.0%; Pred. No. 4e-25;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSNRIILNKALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43
Db 1 MLSNRIILNKALRKAHTSMVRNFRYKPVQSQVQLKPRDLC 43

RESULT 2
US-08-765-244-1
; Sequence 1, Application US/08765244

Publication No. US20010008771A1

GENERAL INFORMATION:
 APPLICANT: Seibel, Peter
 APPLICANT: Seibel, Andrea
 TITLE OF INVENTION: CHEMICAL PEPTIDE-NUCLEIC ACID
 TITLE OF INVENTION: FRAGMENT, PROCESS FOR PRODUCING THE SAME AND ITS USES FOR
 TITLE OF INVENTION: APPROPRIATELY INTRODUCING NUCLEIC ACIDS INTO CELL ORGANELLES
 TITLE OF INVENTION: AND CELLS
 FILE REFERENCE: 8484-0018-999
 CURRENT APPLICATION NUMBER: US/08/765,244
 CURRENT FILING DATE: 1997-10-30
 PRIOR APPLICATION NUMBER: PCT/DE95/00775
 PRIOR FILING DATE: 1995-06-11
 PRIOR APPLICATION NUMBER: DE P 44 21 079.5
 PRIOR FILING DATE: 1994-06-16
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 41
 TYPE: PRT
 ORGANISM: Rattus rattus
 FEATURE:
 US-08-765-244-1

Query Match 90.8%; Score 197; DB 8; Length 41;
 Best Local Similarity 95.3%; Pred. No. 3.7e-22;
 Matches 41; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSOVLKPRDLC 43
 DB 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSOVLKPRDLC 41

RESULT 3

US-10-061-395-48
 Sequence 48; Application US/10061395
 Publication No. US20020192675A1
 GENERAL INFORMATION:
 APPLICANT: Zauderer, Maurice
 APPLICANT: Smith, Ernest S.
 TITLE OF INVENTION: Methods of Identifying Regulator Molecules
 FILE REFERENCE: 1821.0080003
 CURRENT APPLICATION NUMBER: US/10/061,395
 CURRENT FILING DATE: 2002-02-04
 PRIOR APPLICATION NUMBER: 60/271,423
 PRIOR FILING DATE: 2001-02-27
 PRIOR APPLICATION NUMBER: 60/265,880
 PRIOR FILING DATE: 2001-02-05
 PRIOR APPLICATION NUMBER: 60/265,589
 PRIOR FILING DATE: 2001-02-02
 NUMBER OF SEQ ID NOS: 116
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 48
 LENGTH: 32
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Heterologous signal sequence for the mitochondrial matrix
 NAME/KEY: MISC_FEATURE
 LOCATION: (7)..(8)
 OTHER INFORMATION: May be any amino acid
 NAME/KEY: MISC_FEATURE
 LOCATION: (32)..(32)
 OTHER INFORMATION: May be any amino acid
 US-10-061-395-48

Query Match 42.4%; Score 92; DB 14; Length 32;
 Best Local Similarity 61.3%; Pred. No. 1.3e-06;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPV 31
 DB 1 MLFNLRXXLNNAAFRGHNFVNRFCGQPL 31

RESULT 4

US-10-052-942-54
 Sequence 54; Application US/10052942
 Publication No. US20030104402A1
 GENERAL INFORMATION:
 APPLICANT: Zauderer, Maurice
 APPLICANT: Smith, Ernest
 APPLICANT: Wei, Chungwen
 TITLE OF INVENTION: Methods of Producing or Identifying Intrabodies in Eukaryotic C
 FILE REFERENCE: 1821.0090004
 CURRENT APPLICATION NUMBER: US/10/052,942
 CURRENT FILING DATE: 2002-01-23
 PRIOR APPLICATION NUMBER: 60/298,095
 PRIOR FILING DATE: 2001-06-15
 PRIOR APPLICATION NUMBER: 60/271,422
 PRIOR FILING DATE: 2001-02-27
 PRIOR APPLICATION NUMBER: 60/263,200
 PRIOR FILING DATE: 2001-01-24
 PRIOR APPLICATION NUMBER: 60/263,225
 PRIOR FILING DATE: 2001-01-23
 NUMBER OF SEQ ID NOS: 154
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 54
 LENGTH: 32
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: signal sequence
 NAME/KEY: UNSURE
 LOCATION: (7)..(8)
 OTHER INFORMATION: Xaa may represent any amino acid
 NAME/KEY: UNSURE
 LOCATION: (32)..(32)
 OTHER INFORMATION: Xaa may represent any amino acid
 US-10-052-942-54

Query Match 42.4%; Score 92; DB 15; Length 32;
 Best Local Similarity 61.3%; Pred. No. 1.3e-06;
 Matches 19; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPV 31
 DB 1 MLFNLRXXLNNAAFRGHNFVNRFCGQPL 31

RESULT 5

US-09-998-027-3
 Sequence 3; Application US/09998027
 Publication No. US20030093819A1
 GENERAL INFORMATION:
 APPLICANT: D'Andrea et al.
 TITLE OF INVENTION: Methods and Compositions for the
 TITLE OF INVENTION: Diagnosis and Treatment of Cancers Associated with Defective
 TITLE OF INVENTION: DNA Repair Mechanisms
 FILE REFERENCE: 2486/101
 CURRENT APPLICATION NUMBER: US/09/998,027
 CURRENT FILING DATE: 2001-11-02
 NUMBER OF SEQ ID NOS: 191
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 1286
 TYPE: PRT
 ORGANISM: A. thaliana
 FEATURE:
 NAME/KEY: PEPTIDE
 LOCATION: (1)...(1286)
 OTHER INFORMATION: Plantfancd2
 US-09-998-027-3

Query Match 28.1%; Score 61; DB 11; Length 1286;
 Best Local Similarity 38.2%; Pred. No. 4.7;

```

Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 3 SNLRLLNKAALRKAHTSMVRNFRYKGPVQSQVQ 36
Db 107 SNLRMLSSSTTRKDESLVRLNLLVSPQLDIQ 140

RESULT 6
US-10-165-099-3
; Sequence 3, Application US/10165099
; Publication No. US20030189326A1
; GENERAL INFORMATION:
; APPLICANT: D'Andrea, Alan
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS OF CANCER SUSCEPTIBILITY
; FILE REFERENCE: 7032/2055
; CURRENT APPLICATION NUMBER: US/10/165,099
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 09/998,027
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/245,756
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 352
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-165-099-3

Query Match 28.1%; Score 61; DB 12; Length 1286;
Best Local Similarity 38.2%; Pred. No. 4.7;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;
QY 3 SNLRLLNKAALRKAHTSMVRNFRYKGPVQSQVQ 36
Db 107 SNLRMLSSSTTRKDESLVRLNLLVSPQLDIQ 140

RESULT 7
US-10-353-856-8
; Sequence 8, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawacka, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Bas
; FILE REFERENCE: 800.032U51
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2212
; TYPE: PRT
; ORGANISM: Reston Ebola virus
US-10-353-856-8

Query Match 23.3%; Score 50.5; DB 12; Length 2212;
Best Local Similarity 25.6%; Pred. No. 3.3e+02;
Matches 11; Conservative 12; Mismatches 17; Indels 3; Gaps 1;
QY 1 MSLNRLNKAALRKAHTSMVRNFRYKGPVQSQVQLKPRDLQ 43
Db 327 VINDRLNRLNKAALRKAHTSMVRNFRYKGPVQSQVQLKPRDLQ 366

RESULT 8
US-10-156-761-7915

```

```

; Sequence 7915, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7915
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-7915

Query Match 23.0%; Score 50; DB 15; Length 414;
Best Local Similarity 34.3%; Pred. No. 53;
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
QY 6 RLLNKAALRKAHTSMVRNFRYKGPVQSQVQLKPR 40
Db 294 RLQNTGAKPAFARLTRAWRYGGTGTITVAEQPR 328

RESULT 9
US-10-032-585-7102
; Sequence 7102, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7102
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7102

Query Match 23.0%; Score 50; DB 12; Length 564;
Best Local Similarity 29.1%; Pred. No. 77;
Matches 16; Conservative 6; Mismatches 17; Indels 16; Gaps 1;
QY 2 LSNRLNLLNKAALR-----KAHTSMVRNFRYKGPVQSQVQLKPR 40
Db 424 LSDLLELLVIAVRVWIEKPELOQTINFLAYTEYQEMVKNFNTGNAIVSSTLDSR 478

RESULT 10
US-09-950-046A-2
; Sequence 2, Application US/09950046A
; Patent No. US20020151701A1
; GENERAL INFORMATION:
; APPLICANT: Kwiatkowski, David J.
; Sampson, Julian R.
; Povey, Sue
; van Slegtenhorst, Marjon

```

```

; Halley, Dicky
; TITLE OF INVENTION: Compositions and Methods Based Upon the Tuberculous
; ; Sclerosis-1 (TSC-1) Gene and Gene Product
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/950,046A
; FILING DATE: 12-Sep-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BRI331/42002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 639-6585
; TELEFAX: (202) 639-6604
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-950-046A-2

```

```

Query Match 22.8%; Score 49.5; DB 10; Length 1164;
Best Local Similarity 29.8%; Pred. No. 2.2e+02;
Matches 14; Conservative 9; Mismatches 19; Indels 5; Gaps 1;

```

```

QY 1 MLSNRLINKAALRAKHAHSW-----RNFYKPKVQSQVLKPRDL 42
Db 806 MIAELRIELKANKVCHTELLLSQVLSNSESQQQMEFLNROL 852

```

```

RESULT 11
US-09-864-761-46024
; Sequence 46024, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46024
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: MAP TO AC011469.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: SWISSPROT HIT: P16960, EVALUE 3.00e-27
US-09-864-761-46024

```

```

Query Match 22.4%; Score 48.5; DB 9; Length 75;
Best Local Similarity 40.0%; Pred. No. 12;
Matches 14; Conservative 3; Mismatches 11; Indels 7; Gaps 1;

```

```

QY 11 KAALRKAHTS-----MVRNFRYCKPVQSQVLK 38
Db 11 KAGLSFFESASEDIKRVNLRGLCKVQARTQVK 45

```

```

RESULT 12
US-09-815-242-4869
; Sequence 4869, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578

```

```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4869
/ LENGTH: 285
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
/ US-09-815-242-4869

Query Match      22.4%; Score 48.5; DB 9; Length 285;
Best Local Similarity 34.9%; Pred. No. 57;
Matches 15; Conservative 6; Mismatches 17; Indels 5; Gaps 2;

QY 5 LRLLNK-AALRKAHTSMVR-----NFRYKPVOSQVQLKPRDL 42
Db 147 LRRLNKLATRETHFRVEMVSGFKYGLPIDADIVMDVRF 189

RESULT 13
US-09-815-242-10798
/ Sequence 10798, Application US/09815242
/ Patent No. US20020061569A1
/ GENERAL INFORMATION:
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari L.
/ APPLICANT: Zyskind, Judith W.
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John D.
/ APPLICANT: Carr, Grant J.
/ APPLICANT: Yamamoto, Robert T.
/ APPLICANT: Xu, H. Howard
/ TITLE OF INVENTION: Identification of Essential Genes in
/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ PRIOR FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10798
/ LENGTH: 334
/ TYPE: PRT
/ ORGANISM: Enterococcus faecalis
/ US-09-815-242-10798

Query Match      22.4%; Score 48.5; DB 9; Length 334;
Best Local Similarity 34.9%; Pred. No. 69;
Matches 15; Conservative 6; Mismatches 17; Indels 5; Gaps 2;

QY 5 LRLLNK-AALRKAHTSMVR-----NFRYKPVOSQVQLKPRDL 42
Db 189 LRRLNKLATRETHFRVEMVSGFKYGLPIDADIVMDVRF 231

RESULT 14
US-09-738-626-3622
/ Sequence 3622, Application US/09738626
/ Publication No. US20020197605A1
/ GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIKO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIALI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAKO
/ APPLICANT: SENOH, AKIHITO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: Patent in ver. 3.0
/ SEQ ID NO 3622
/ LENGTH: 684
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
/ US-09-738-626-3622

Query Match      22.4%; Score 48.5; DB 10; Length 684;
Best Local Similarity 30.2%; Pred. No. 1.6e+02;
Matches 13; Conservative 10; Mismatches 15; Indels 5; Gaps 2;

QY 3 SNRLILNKAAAL---RKAHTSMVRNRYKPVOSQVQLKPRDL 42
Db 76 NSLHLRNGSDILRNHHCWVFNF--NRPDPVRLGPRDI 116

RESULT 15
US-09-764-891-4156
/ Sequence 4156, Application US/09764891
/ Publication No. US20030077808A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ FILE REFERENCE: PC006
/ CURRENT APPLICATION NUMBER: US/09/764,891
/ CURRENT FILING DATE: 2001-01-17
/ PRIOR application data removed - consult PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 10231
/ SOFTWARE: Patent in Ver. 2.0
/ SEQ ID NO 4156
/ LENGTH: 235
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (208)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (221)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (223)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (228)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
```

```
; LOCATION: (231)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (234)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4156
```

```
Query Match          22.1%; Score 48; DB 11; Length 235;
Best Local Similarity 37.0%; Pred. No. 54;
Matches 10; Conservative 7; Mismatches 10; Indels 0; Gaps 0;
```

```
QY      9 LNKALRKKAHTSMVRNFRYKPKVQSQV 35
      | : : ||| : ||| : ||| :
Db      63 LGQGLRKGRSSMRKNGSLQRP LQSGI 89
```

```
Search completed: December 3, 2003, 14:35:35
Job time : 30 secs
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:30:02 ; Search time 21 Seconds
(without alignments)
196.917 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217
Sequence: 1 MLSNRLILNKALKAHTS.....NFRYKPYQSQVQLKPRDLC 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:**

1: Pirl:**
2: Pirl:**
3: Pirl:**
4: Pirl:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199	91.7	354	1 OWRT	ornithine carbamoyl
2	164	75.6	354	1 OWMS	ornithine carbamoyl
3	141	65.0	354	1 OWHU	ornithine carbamoyl
4	119	54.8	41	2 I52779	ornithine transcar
5	61	28.1	1286	2 B71413	hypothetical prote
6	60	27.6	351	2 F90409	GTP binding conser
7	57	26.3	354	2 JE0309	ornithine carbamoy
8	55.5	25.6	454	2 A86345	Fl6P4.13 protein -
9	55	25.3	350	2 A48421	ornithine transcar
10	54	24.9	300	2 T32702	hypothetical prote
11	54	24.9	457	2 A09550	oxygen-independent
12	54	24.9	1225	2 A56514	chromokinesin - ch
13	53.5	24.7	423	2 I40646	sensor-like protei
14	53	24.4	435	2 S31290	cyclin B5 - yeast
15	53	24.4	593	2 T47000	nitrate/nitrite se
16	53	24.4	593	2 AB0239	nitrate/nitrite se
17	52.5	24.2	185	2 B81708	translation elonga
18	52	24.0	213	2 B83182	probable two-compo
19	52	24.0	335	2 T33457	hypothetical prote
20	52	24.0	598	1 RGEONX	nitrate/nitrite se
21	52	24.0	598	2 F85702	nitrate/nitrite se
22	52	24.0	598	2 G90844	nitrate/nitrite se
23	51.5	23.7	280	2 G70233	hypothetical prote
24	51	23.5	506	2 T50211	WD-repeat protein
25	51	23.5	598	2 AF0648	nitrate/nitrite se
26	51	23.5	633	2 S76749	hypothetical prote
27	51	23.5	705	2 S54521	probable membrane
28	51	23.5	1779	2 T23130	hypothetical prote
29	50.5	23.3	138	2 F81900	hypothetical prote

ALIGNMENTS

RESULT 1

OWRT

ornithine carbamoyltransferase (EC 2.1.3.3) precursor - rat

N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 28-Feb-1986 #sequence revision 28-Feb-1986 #text change 11-Jun-1999

C:Accession: A00563; A28042; A23090; S02466; I52976; I67609; I53457

R:Takiguchi, M.; Miura, S.; Mori, M.; Tatibana, M.; Nagata, S.; Kaziro, Y.

Proc. Natl. Acad. Sci. U.S.A. 81, 7412-7416, 1984

A>Title: Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyl

A:Reference number: A00563; MUID:85063800; PMID:6095294

A:Accession: A00563

A:Molecule type: mRNA

A:Residues: 1-354 <TAK1>

A:Cross-references: GB:K03040; NID:G205873; PID:AAA41768.1; PID:G205874

R:Takiguchi, M.; Murakami, T.; Miura, S.; Mori, M.

Proc. Natl. Acad. Sci. U.S.A. 84, 6136-6140, 1987

A>Title: Structure of the rat ornithine carbamoyltransferase gene, a large, X chromos

A:Reference number: A28042; MUID:87317609; PMID:3476935

A:Accession: A28042

A:Molecule type: DNA

A:Residues: 1-354 <TAK2>

A:Cross-references: GB:M16933; GB:J02957; NID:G205884; PID:AAA41769.1; PID:G205886

R:Kraus, J.P.; Hodges, P.E.; Williamson, C.L.; Horwich, A.L.; Kalousek, F.; Williams,

Nucleic Acids Res. 13, 943-952, 1985

A>Title: A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase

A:Reference number: A23090; MUID:85215524; PMID:3839075

A:Accession: A23090

A:Molecule type: mRNA

A:Residues: 1-38, 'P', '40-240, 'S', '242-354 <KRA>

R:Aoki, Y.; Sunaga, H.; Suzuki, K.T.

Biochem. J. 250, 735-742, 1988

A>Title: A cadmium-binding protein in rat liver identified as ornithine carbamoyltrans

A:Reference number: S02466; MUID:88268748; PMID:3390141

A:Accession: S02466

A:Molecule type: protein

A:Residues: 33-56;293-302;307-317;322-329 <AOK>

R:McIntyre, P.; Graf, L.; Mercer, J.F.B.; Wake, S.A.; Hudson, P.J.; Hoogenraad, N.

DNA 4, 147-156, 1985

A>Title: The primary structure of the imported mitochondrial protein, ornithine trans

A:Reference number: I52976; MUID:85203360; PMID:3838931

A:Accession: I52976

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-354 <RES>

A:Cross-references: GB:M11266; NID:G205871; PID:AAA41767.1; PID:G205872

R:McIntyre, P.; Graf, L.; Mercer, J.; Peterson, G.; Hudson, P.J.; Hoogenraad, N.

FEBS Lett. 177, 41-46, 1984

A>Title: A highly basic N-terminal extension of the mitochondrial matrix enzyme ornith

A:Reference number: I53457; MUID:85051832; PMID:6548714

A:Accession: I67609

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-102 <RE2>

A:Cross-references: EMBL:X01178; NID:G56802; PIDN:CAA25618.1; PID:G56803

A:Accession: I53457

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-43, 'N', 45-99, 'R', 101-102 <RE3>

A:Cross-references: GB:X03041; NID:G205989; PIDN:AAA41771.1; PID:G205990

C:Genetics:

A:Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3; 335/3

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

C:Keywords: arginine biosynthesis; homotrimer; mitochondrion; transferase; urea cycle

F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:33-354/Product: ornithine carbamoyltransferase #status predicted <MAT>

F:40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match

Best Local Similarity 91.7%; Score 199; DB 1; Length 354;

Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

DB 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

RESULT 2

OMW

ornithine carbamoyltransferase (EC 2.1.3.3) precursor - mouse

N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1993 #sequence revision 13-Mar-1997 #text_change 11-Jun-1999

A:Accession: A43609; S03407; I55252

A:Cross-references: GB:M17030; NID:G200162; PIDN:AAA39865.1; PID:G200163

A:Reference number: 1593-1501, 1988

A:Title: The genetic structure of mouse ornithine transcarbamylase.

A:Reference number: S03407; MUID:88157717; PMID:2831503

A:Accession: S03407

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-194, 'R', 196-335 <SCH>

A:Cross-references: EMBL:X07092

R:Veres, G.; Craig, W.J.; Caskey, C.T.

J. Biol. Chem. 261, 7588-7591, 1986

A:Title: The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequence

A:Reference number: I55252; MUID:86224037; PMID:3011788

A:Accession: I55252

A:Status: translation not shown; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-19, 'LLWDFIF' <RES>

A:Cross-references: GB:M12716; NID:G200160; PIDN:AAA39864.1; PID:G554248

A:Note: the end of this sequence is near the boundary of the cloned region and may be an artifact

C:Genetics:

A:Gene: ONC

A:Map position: X

A:Introns: 26/2; 72/3; 100/1; 129/2; 180/3; 221/3; 239/3; 289/3

C:Superfamily: ornithine carbamoyltransferase; aspartate/ornithine carbamoyltransferase

C:Keywords: mitochondrion; transferase; urea cycle

F:1-32/Domain: transit peptide (mitochondrion) #status predicted <TNP>

F:33-335/Product: ornithine carbamoyltransferase #status predicted <MAT>

F:40-342/Domain: aspartate/ornithine carbamoyltransferase homology <ACT>

Query Match

Best Local Similarity 75.6%; Score 164; DB 1; Length 354;

Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

DB 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

DB 1 MLSNRLINKAALRKAHTSMVNFYRGKPVQSQVQLKGRDL 42

RESULT 3

OMW

ornithine carbamoyltransferase (EC 2.1.3.3) precursor - human

N:Alternate names: citrulline phosphorylase; ornithine transcarbamylase

C:Species: Homo sapiens (man)

C>Date: 28-Feb-1986 #sequence revision 31-Mar-1993 #text_change 16-Jun-2000

A:Accession: A41444; B41444; A00562; I38078; JC4672; I54377

R:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.

J. Biochem. 103, 302-308, 1988

A:Title: Structure of the human ornithine transcarbamylase gene.

A:Reference number: A41444; MUID:88227905; PMID:2836378

A:Accession: A41444

A:Molecule type: DNA

A:Residues: 1-354 <CHAT>

A:Cross-references: GB:D00230; NID:G219957; PIDN:BAA00161.1; PID:G219959

A:Accession: B41444

A:Molecule type: mRNA

A:Residues: 1-354 <HA2>

A:Cross-references: GB:D00230; NID:G219957; PIDN:BAA00161.1; PID:G219959

R:Horwich, A.L.; Fenton, W.A.; Williams, K.R.; Kalousek, F.; Kraus, J.P.; Doolittle, R.

Science 224, 1068-1074, 1984

A:Title: Structure and expression of a complementary DNA for the nuclear coded precursor

A:Reference number: A00562; MUID:84196410; PMID:6372096

A:Accession: A00562

A:Molecule type: mRNA

A:Residues: 1-100, 'F', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <HOR>

A:Cross-references: GB:D00230

R:Hata, A.; Tsuzuki, T.; Shimada, K.; Takiguchi, M.; Mori, M.; Matsuda, I.

J. Biochem. 100, 717-725, 1986

A:Title: Isolation and characterization of the human ornithine transcarbamylase gene.

A:Reference number: I38078; MUID:87057134; PMID:3782067

A:Accession: I38078

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-26 <RES>

A:Cross-references: EMBL:X04443; NID:G35162; PIDN:CAA28039.1; PID:G35163

R:Wheeler, V.C.; Prodomou, C.P.; Pearl, L.H.; Williamson, R.; Coutelle, C.

Gene 169, 251-255, 1996

A:Title: Synthesis of a modified gene encoding human ornithine transcarbamylase for ex

pression in

A:Reference number: JC4672; MUID:96194812; PMID:8647457

A:Accession: JC4672

A:Molecule type: DNA

A:Residues: 'M', 33-100, 'F', 102-110, 'P', 112-192, 'CF', 195-269, 'R', 271-354 <WHE>

A:Note: this report represents a synthetic gene designed for expression in (rather than

R:Horwich, A.L.; Kalousek, F.; Rosenberg, L.E.

Proc. Natl. Acad. Sci. U.S.A. 82, 4930-4933, 1985

A:Title: Arginine in the leader peptide is required for both import and proteolytic cleavage

A:Reference number: I59039; MUID:85270440; PMID:3895227

A:Accession: I59039

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-36 <RE2>

A:Cross-references: GB:M1235; NID:G189408; PIDN:AAA59976.1; PID:G189409

R:Gilbert-Dussardier, B.; Rabier, D.; Strautnieks, S.; Segues, B.; Bonnefont, J.P.; Mu

Hum. Mol. Genet. 3, 831-832, 1994

A:Title: A novel arginine (245) to glutamine change in exon 8 of the ornithine carbamo

c pattern.

A:Reference number: I54377; MUID:94362689; PMID:8081373

A:Accession: I54377

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 269-276, 'Q', 278-289 <RE3>

A:Cross-references: GB:S73640; NID:G688001; PIDN:AAB31859.1; PID:G688002

A:Note: this sequence represents a disease defect in ornithine carbamoyltransferase

C:Comment: The active enzyme is a dimer of identical chains with one tightly bound zin

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:17:57 ; Search time 11 Seconds
(without alignments)
183.832 Million cell updates/sec

Title: US-08-765-244-22

Perfect score: 217

Sequence: 1 MLSNRLILNKAALRKAHTS.....NFRYKPVQSQVQLKPRDLC 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt 41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	91.7	354	1 OTC RAT	P00481 rattus norv
2	164	75.6	354	1 OTC MOUSE	P11725 mus musculus
3	141	65.0	354	1 OTC HUMAN	P00480 homo sapien
4	59	27.2	597	1 IF2P METAC	Q8qkl5 methanosarc
5	56	25.8	591	1 IF2P METWA	Q9pu78 methanosarc
6	55	25.3	350	1 OTC RANCA	P31326 rana catesb
7	54	24.9	236	1 RR2 LOTJA	Q9bbs6 lotus japon
8	54	24.9	457	1 HEMN SALTY	P37129 salmonella
9	54	24.9	1225	1 OTC CHICK	Q90640 gallus gall
10	53	24.4	328	1 OTC FIG	O19072 sus scrofa
11	53	24.4	435	1 C8S5 YEAST	P30283 saccharomyc
12	52.5	24.2	185	1 EFP1 CHLMU	Q9pkx6 chlamydia m
13	52	24.0	598	1 NARX ECOLI	P10956 escherichia
14	51	23.5	506	1 POFB SCHPO	Q09855 schizosacch
15	51	23.5	705	1 YM37 YEAST	O03824 saccharomyc
16	51	23.5	1169	1 C3DA BACTP	O05014 bacillus th
17	50.5	23.3	429	1 TRB2 SULSO	Q97tx6 sulfolobus
18	50	23.0	330	1 LAFU VIBPA	Q03478 vibrio para
19	50	23.0	622	1 SR68 CANFA	Q00004 canis famil
20	50	23.0	662	1 PRCA ANASP	Q59149 anabaena sp
21	49.5	22.8	864	1 RA50 SULSO	Q97wh0 sulfolobus
22	49.5	22.8	1163	1 TSCL1 RAT	Q92136 rattus norv
23	49.5	22.8	1164	1 TSCL1 HUMAN	Q92574 homo sapien
24	49	22.6	457	1 HEMN ECOLI	P32131 escherichia
25	49	22.6	1219	1 MYS5 YEAST	Q04439 saccharomyc
26	49	22.6	1345	1 YH00 YEAST	P38800 saccharomyc
27	48.5	22.4	5035	1 RYR1 PIG	P16960 sus scrofa
28	48.5	22.4	5037	1 RYR1 RABIT	P11716 cryptolagus
29	48.5	22.4	5038	1 RYR1 HUMAN	P21817 homo sapien
30	48	22.1	88	1 R37A SCHPO	P59289 schizosacch
31	48	22.1	578	1 SYR BUCBP	P59483 buchnera ap
32	48	22.1	647	1 PRCA ANAVA	P23916 anabaena va
33	48	22.1	1025	1 BGAL_KLULA	P00723 kluyveromyc

34	47.5	21.9	333	1 Y376 BUCAP	Q8k951 buchnera ap
35	47.5	21.9	1473	1 OVOS_CHICK	P20740 gallus galli
36	47.5	21.9	2210	1 RRPL_EBOSM	Q66802 ebola virus
37	47	21.7	90	1 R37B SCHPO	P05733 schizosacch
38	47	21.7	260	1 PSMA PYRAB	Q9V122 pyrococcus
39	47	21.7	260	1 PSMA PYRHO	O95219 pyrococcus
40	47	21.7	421	1 TES HUMAN	Q9UG18 homo sapien
41	47	21.7	423	1 TES_MOUSE	P47226 mus musculus
42	47	21.7	1098	1 RPOB_MAIZE	P10581 zea mays (m
43	46.5	21.4	410	1 T801_PSESH	P24607 pseudomonas
44	46.5	21.4	421	1 VCOM_ADECC	Q65952 canine aden
45	46.5	21.4	421	1 VCOM_ADECR	Q96685 canine aden

ALIGNMENTS

RESULT 1
OTC_RAT
ID OTC RAT STANDARD; PRT; 354 AA.
AC P00481; Q63407;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
DE (OTCase) (Ornithine transcarbamylase).
DE OTC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=85063800; PubMed=6095294;
RA Takiguchi M., Miura S., Mori M., Tatibana M., Nagata S., Kaziro Y.;
RT "Molecular cloning and nucleotide sequence of cDNA for rat ornithine carbamoyltransferase precursor."
RL Proc. Natl. Acad. Sci. U.S.A. 81:7412-7416(1984).
RN [2]
SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Liver;
MEDLINE=87317609; PubMed=3476935;
RA Takiguchi M., Murakami T., Miura S., Mori M.;
RT "Structure of the rat ornithine carbamoyltransferase gene, a large, X chromosome-linked gene with an atypical promoter."
RL Proc. Natl. Acad. Sci. U.S.A. 84:6136-6140(1987).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=85215524; PubMed=3839075;
RA Kraus J.P., Hodges P.E., Williamson C.L., Horwich A.L., Kalousek F., Williams K.R., Rosenberg L.E.;
RT "A cDNA clone for the precursor of rat mitochondrial ornithine transcarbamylase: comparison of rat and human leader sequences and conservation of catalytic sites."
RL Nucleic Acids Res. 13:943-952(1985).
RN [4]
SEQUENCE FROM N.A.
RX MEDLINE=85203360; PubMed=3838931;
RA McIntyre P., Graf L., Mercer J.F.B., Wake S.A., Hudson P.J., Hoogenraad N.;
RT "The primary structure of the imported mitochondrial protein, ornithine transcarbamylase from rat liver: mRNA levels during ontogeny."
RL DNA 4:147-156(1985).
RN [5]
SEQUENCE OF 33-56; 293-302; 307-317 AND 321-329.
RC TISSUE=Liver;
RX MEDLINE=88268748; PubMed=3390141;
RA Aoki Y., Sunaga H., Suzuki K.T.;
RT "A cadmium-binding protein in rat liver identified as ornithine carbamoyltransferase."
RL Biochem. J. 250:735-742(1988).
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate

```
CC + L-citrulline.
CC -!- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
CC -!- SUBUNIT: Homotrimer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K03040; AAA41768.1; -
DR EMBL; M16933; AAA41769.1; -
DR EMBL; M16924; AAA41769.1; JOINED.
DR EMBL; M16925; AAA41769.1; JOINED.
DR EMBL; M16926; AAA41769.1; JOINED.
DR EMBL; M16928; AAA41769.1; JOINED.
DR EMBL; M16929; AAA41769.1; JOINED.
DR EMBL; M16930; AAA41769.1; JOINED.
DR EMBL; M16932; AAA41769.1; JOINED.
DR EMBL; X01976; CAA26007.1; -
DR EMBL; K00001; AAA41772.1; -
DR EMBL; M11266; AAA41767.1; -
DR HSP; A00563; OWRT.
DR HSP; P00480; 10TH.
DR InterPro; IPR006130; Asp/Orn Cotransf.
DR InterPro; IPR002292; Orn_carbtransf.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace.N; 1.
DR PRINTS; PR00100; AOTCACE.
DR TIGRFS; TIGR00658; orn_carb_tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
KW Transit peptide.
FT CHAIN 1 32 MITOCHONDRION.
FT ACT SITE 263 354 ORNITHINE CARBAMOYLTRANSFERASE.
FT ACT SITE 303 303 BY SIMILARITY.
FT ACT SITE 39 39 G -> P (IN REF. 3; AAA41772).
FT CONFLICT 241 241 G -> S (IN REF. 3).
SQ SEQUENCE 354 AA; 39886 MW; 156B511AF7063F0C CRC64;

Query Match 91.7%; Score 199; DB 1; Length 354;
Best Local Similarity 97.6%; Pred. No. 9.2e-21;
Matches 41; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSNRLILNKAALKAHKAHTSMVNFYRGKVFQSQVQLKPRDL 42
Dd 1 MLSNRLILNKAALKAHKAHTSMVNFYRGKVFQSQVQLKGRDL 42

RESULT 2
OTC_MOUSE STANDARD; PRT; 354 AA.
AC P11725;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
DE (Otcase) (Ornithine transcarbamylase).
GN OTC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
```

```
RX MEDLINE=87263407; PubMed=3603027;
RA Veres G., Gibbs R.A., Scherer S.E., Caskey C.T.;
RT "The molecular basis of the sparse fur mouse mutation.";
RL Science 237:415-417(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=8815717; PubMed=2831503;
RA Scherer S.E., Veres G., Caskey C.T.;
RT "The genetic structure of mouse ornithine transcarbamylase.";
RL Nucleic Acids Res. 16:1593-1601(1988).
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=86224037; PubMed=3011788;
RA Veres G., Craigen W.J., Caskey C.T.;
RT "The 5' flanking region of the ornithine transcarbamylase gene
contains DNA sequences regulating tissue-specific expression.";
RL J. Biol. Chem. 261:7588-7591(1986).
CC -!- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
+ L-citrulline.
CC -!- PATHWAY: SECOND STEP IN UREA CYCLE, ARGININE BIOSYNTHESIS.
CC -!- SUBUNIT: Homotrimer.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- DISEASE: SPARSE FUR (SPF) MOUSE HAVE AN OTC WITH AN OVERALL
DECREASE IN ACTIVITY, AND ALTERED SUBSTRATE AFFINITY.
CC -!- SIMILARITY: BELONGS TO THE ATCASES/OTCASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17030; AAA39865.1; -
DR EMBL; M13716; AAA39864.1; ALT_SEQ.
DR EMBL; X07092; CAA30121.1; -
DR EMBL; X07093; CAA30121.1; JOINED.
DR EMBL; X07094; CAA30121.1; JOINED.
DR EMBL; X07095; CAA30121.1; JOINED.
DR EMBL; X07096; CAA30121.1; JOINED.
DR EMBL; X07097; CAA30121.1; JOINED.
DR EMBL; X07098; CAA30121.1; JOINED.
DR EMBL; X07099; CAA30121.1; JOINED.
DR EMBL; X07100; CAA30121.1; JOINED.
DR PIR; A43609; OWMS.
DR HSP; P00480; 10TH.
DR SWISS-2DPAGE; P11725; MOUSE.
DR MGD; MGI:97448; Otc.
DR InterPro; IPR006130; Asp/Orn Cotransf.
DR InterPro; IPR002292; Orn_carbtransf.
DR InterPro; IPR006131; OTCace_O.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF00185; OTCace; 1.
DR Pfam; PF02729; OTCace.N; 1.
DR PRINTS; PR00100; AOTCACE.
DR TIGRFS; TIGR00658; orn_carb_tr; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR Arginine biosynthesis; Urea cycle; Transferase; Mitochondrion;
KW Transit peptide.
FT CHAIN 1 32 MITOCHONDRION.
FT ACT SITE 263 354 ORNITHINE CARBAMOYLTRANSFERASE.
FT ACT SITE 303 303 BY SIMILARITY.
FT ACT SITE 39 39 H -> N (IN SPARSE FUR MOUSE).
FT VARIANT 117 117 G -> R (IN REF. 2).
FT CONFLICT 195 195 33BBESD1B8AA196 CRC64;
SQ SEQUENCE 354 AA; 39765 MW; 33BBESD1B8AA196 CRC64;

Query Match 75.6%; Score 164; DB 1; Length 354;
Best Local Similarity 83.3%; Pred. No. 8.8e-16;
Matches 35; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

QY 1 MSLNRLILNKAAALRKATSMVRNPKYQVQSQVQLKPRDL 42
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 MSLNRLILNNAALRKCHTSVVRHFWCKGPKYQVQSQVQLKGRDL 42

RESULT 3
 OTC HUMAN STANDARD; PRT; 354 AA.
 AC P00480;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3)
 DE (OrnCase) (Ornithine transcarbamylase).
 OTC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=88227905; PubMed=2836378;
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;
 RT "Structure of the human ornithine transcarbamylase gene.";
 RL J. Biochem. 103:302-308 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=84196410; PubMed=6372096;
 RA Horwich A.L., Fenton W.A., Williams K.R., Kalousek F., Kraus J.P.,
 RA Doolittle R.F., Konigsberg W., Rosenberg L.E.;
 RT "Structure and expression of a complementary DNA for the nuclear
 RT coded precursor of human mitochondrial ornithine transcarbamylase.";
 RL Science 224:1068-1074 (1984).
 RN [3]
 RP SEQUENCE OF 1-36 FROM N.A.
 RX MEDLINE=85270440; PubMed=3895227;
 RA Horwich A.L., Kalousek F., Rosenberg L.E.;
 RT "Arginine in the leader peptide is required for both import and
 RT proteolytic cleavage of a mitochondrial precursor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4930-4933 (1985).
 RN [4]
 RP SEQUENCE OF 1-26 FROM N.A.
 RX MEDLINE=87057134; PubMed=3782067;
 RA Hata A., Tsuzuki T., Shimada K., Takiguchi M., Mori M., Matsuda I.;
 RT "Isolation and characterization of the human ornithine
 RT transcarbamylase gene: structure of the 5'-end region.";
 RL J. Biochem. 100:717-725 (1986).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RX MEDLINE=99069419; PubMed=9852088;
 RA Shi D., Morizono H., Ha Y., Aoyagi M., Tuchman M., Allewell N.M.;
 RT "1.85-A resolution crystal structure of human ornithine
 RT transcarbamoylase complexed with N-phosphonacetyl-L-ornithine.
 RT Catalytic mechanism and correlation with inherited deficiency.";
 RL J. Biol. Chem. 273:34247-34254 (1998).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=20274073; PubMed=10813810;
 RA Shi D., Morizono H., Aoyagi M., Tuchman M., Allewell N.M.;
 RT "Crystal structure of human ornithine transcarbamylase complexed with
 RT carbamoyl phosphate and L-norvaline at 1.9 A resolution.";
 RL Proteins 39:271-277 (2000).
 RN [7]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=93372869; PubMed=8364586;
 RA Tuchman M.;
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase
 RT gene.";
 RL Hum. Mutat. 2:174-178 (1993).
 RN [8]
 RP REVIEW ON VARIANTS.

RX MEDLINE=95353279; PubMed=7627182;
 RA Tuchman M., Plante R.J.;
 RT "Mutations and polymorphisms in the human ornithine transcarbamylase
 RT gene: mutation update addendum.";
 RL Hum. Mutat. 5:293-295 (1995).
 RN [9]
 RP REVIEW ON VARIANTS, AND 3D-STRUCTURE MODELING.
 RX MEDLINE=96091868; PubMed=8544185;
 RA Tuchman M., Morizono H., Reish O., Yuan X., Allewell N.M.;
 RT "The molecular basis of ornithine transcarbamylase deficiency:
 RT modelling the human enzyme and the effects of mutations.";
 RL J. Med. Genet. 32:680-688 (1995).
 RN [10]
 RP VARIANT OTCD GLN-141.
 RX MEDLINE=89008892; PubMed=3170748;
 RA Maddalena A., Spence J.E., O'Brien W.E., Nussbaum R.L.;
 RT "Characterization of point mutations in the same arginine codon in
 RT three unrelated patients with ornithine transcarbamylase
 RT deficiency.";
 RL J. Clin. Invest. 82:1353-1358 (1988).
 RN [11]
 RP VARIANTS OTCD GLN-26; PRO-45 AND GLU-216, AND VARIANT ARG-46.
 RX MEDLINE=89345570; PubMed=2474822;
 RA Grompe M., Muzny D.M., Caskey C.T.;
 RT "Scanning detection of mutations in human ornithine transcarbamylase
 RT by chemical mismatch cleavage.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:5888-5892 (1989).
 RN [12]
 RP VARIANT OTCD TRP-277.
 RX MEDLINE=90269805; PubMed=2347583;
 RA Finkelstein J.E., Francmano C.A., Brusilow S.W., Travstman M.D.;
 RT "Use of denaturing gradient gel electrophoresis for detection of
 RT mutation and prospective diagnosis in late onset ornithine
 RT transcarbamylase deficiency.";
 RL Genomics 7:167-172 (1990).
 RN [13]
 RP VARIANTS OTCD GLN-92 AND LEU-320, AND VARIANT PRO-111.
 RX MEDLINE=91118929; PubMed=1671317;
 RA Grompe M., Caskey C.T., Fenwick R.G. Jr.;
 RT "Improved molecular diagnostics for ornithine transcarbamylase
 RT deficiency.";
 RL Am. J. Hum. Genet. 48:212-222 (1991).
 RN [14]
 RP VARIANT OTCD LEU-225.
 RX MEDLINE=92098086; PubMed=1721894;
 RA Henzen D., Pelet A., Feldman D., Rabier D., Berthelot J.,
 RA Munnich A.;
 RT "Fatal hyperammonemia resulting from a C-to-T mutation at a MspI site
 RT of the ornithine transcarbamylase gene.";
 RL Hum. Genet. 88:153-156 (1991).
 RN [15]
 RP VARIANTS OTCD GLU-79; THR-94; PHE-304 AND ASP-345.
 RX MEDLINE=93126062; PubMed=1480464;
 RA Tuchman M., Holzknecht R.A., Gueron A.B., Berry S.A., Tsai M.Y.;
 RT "Six new mutations in the ornithine transcarbamylase gene detected by
 RT single-strand conformational polymorphism.";
 RL Pediatr. Res. 32:600-604 (1992).
 RN [16]
 RP SEQUENCE OF 269-289 FROM N.A., AND VARIANT OTCD GLN-277.
 RX MEDLINE=94362689; PubMed=8081373;
 RA Gilbert-Dussardier B., Rabier D., Strautnieks S., Segues B.,
 RA Bonnefont J.-P., Munnich A.;
 RT "A novel arginine (245) to glutamine change in exon 8 of the ornithine
 RT carbamoyl transferase gene in two unrelated children presenting with
 RT late onset deficiency and showing the same enzymatic pattern.";
 RL Hum. Mol. Genet. 3:831-832 (1994).
 RN [17]
 RP VARIANT OTCD PRO-140.
 RX MEDLINE=93372936; PubMed=8099056;
 RA Tsai M.Y., Holzknecht R.A., Tuchman M.;
 RT "Single-strand conformational polymorphism and direct sequencing
 RT applied to carrier testing in families with ornithine
 RT transcarbamylase deficiency.";

RESULT 4
IF2P METAC

Ornithine carbamoyltransferase, mitochondrial precursor (EC 2.1.3.3) (OTCase) (Ornithine transcarbamylase).

```

ID  RR2 LOTJA          STANDARD;          PRT;      236 AA.
AC  Q9BES6;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Chloroplast 30S ribosomal protein S2.
GN  RPS2.
OS  Lotus japonicus.
OC  Chloroplast.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
OX  NCBI_TaxID=34305;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Accession MG-20;
RX  MEDLINE=21082929; PubMed=11214967;
RA  Kato T., Kaneko T., Sato S., Nakamura Y., Tabata S.;
RT  "Complete structure of the chloroplast genome of a legume, Lotus
RL  japonicus.";
CC  DNA Res. 7:323-330(2000).
CC  -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AP002983; BAB33197.1; -.
DR  HAMAP; MF_00291; -.
DR  InterPro; IPR001865; Ribosomal S2.
DR  InterPro; IPR005706; S2_bact.org.
DR  Pfam; PF00318; Ribosomal S2_1.
DR  PRINTS; PR00395; RIBOSOMALS2.
DR  TIGRFAMs; TIGR01011; rpsB_bact; 1.
DR  PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR  PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW  Ribosomal protein; Chloroplast.
SQ  SEQUENCE 236 AA; 26982 MW; CE9E238572325586 CRC64;

Query Match      24.9%; Score 54; DB 1; Length 236;
Best Local Similarity 38.5%; Pred. No. 2.6;
Matches 10; Conservative 10; Mismatches 4; Indels 2; Gaps 1;

QY  2 LSNRLILNKA--ALRKAHTSMVRNF 25
Db  211 IASIRILNKLIVFAICEGSHSVIRNF 236

RESULT 8
HEMN SALTY
ID  HEMN SALTY          STANDARD;          PRT;      457 AA.
AC  P37129;
DT  01-OCT-1994 (Rel. 30, Created)
DT  01-OCT-1994 (Rel. 30, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Oxygen-independent coproporphyrinogen III oxidase (EC 1.---)
DE  (Coproporphyrinogenase) (Coprogen oxidase).
GN  HEMN OR STM4004 OR SFY3877 OR T3617.
OS  Salmonella typhimurium, and
OS  Salmonella typhi.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Salmonella.
OX  NCBI_TaxID=602; 601;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  SPECIES=S.typhimurium; STRAIN=LT2;
RX  MEDLINE=94252986; PubMed=8195073;
RA  Xu K., Elliott T.;
RT  "Cloning, DNA sequence, and complementation analysis of the

```

```

RT  Salmonella typhimurium hemN gene encoding a putative
RT  oxygen-independent coproporphyrinogen III oxidase.";
RL  J. Bacteriol. 176:3196-3203(1994).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX  MEDLINE=21534948; PubMed=11677609;
RA  McClelland M., Sanderson K.E., Spieth J., Clifton S.W., LaReille P.,
RA  Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA  Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA  Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA  Waterston R., Wilson R.K.;
RT  "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT  LT2.";
RL  Nature 413:852-856(2001).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  SPECIES=S.typhi; STRAIN=CT18;
RX  MEDLINE=21534947; PubMed=11677608;
RA  Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA  Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA  Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA  Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA  Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA  Whitehead S., Barrall B.G.;
RT  "Complete genome sequence of a multiple drug resistant Salmonella
RT  enterica serovar Typhi CT18.";
RL  Nature 413:848-852(2001).
RN  [4]
RP  SEQUENCE FROM N.A.
RC  SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
RX  MEDLINE=22533367; PubMed=12644504;
RA  Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA  Buxland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT  "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT  and CT18.";
RL  J. Bacteriol. 185:2330-2337(2003).
CC  -!- FUNCTION: ANAEROBIC TRANSFORMATION OF COPROPORPHYRINOGEN-III INTO
CC  PROTOPORPHYRINOGEN-IX.
CC  -!- COFACTOR: REQUIRES MAGNESIUM, ATP AND NAD (OR NADP) FOR ACTIVITY.
CC  -!- PATHWAY: Porphyrin biosynthesis.
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- SIMILARITY: BELONGS TO THE ANAEROBIC COPROPORPHYRINOGEN III
CC  OXIDASE FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; U06779; AAA19690.1; -.
DR  EMBL; AE008887; AAL22843.1; -.
DR  EMBL; AL627280; CAD03096.1; -.
DR  EMBL; AE016846; AA071118.1; -.
DR  StyGene; SG10154; hemN.
DR  InterPro; IPR006638; ELP3.
DR  InterPro; IPR004558; HemN.
DR  Pfam; PF04055; Radical_SAM; 1.
DR  SMART; SM00729; ELP3; 1.
DR  TIGRFAMs; TIGR00538; hemN; 1.
KW  Porphyrin biosynthesis; Oxidoreductase; Magnesium; NAD;
KW  Complete proteome.
SQ  SEQUENCE 457 AA; 52828 MW; 5667B4FE76204DAB CRC64;

Query Match      24.9%; Score 54; DB 1; Length 457;
Best Local Similarity 26.5%; Pred. No. 5.3;
Matches 9; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```



```

CGSS_YEAST
ID CGS5_YEAST STANDARD; PRT; 435 AA.
AC P30283;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S-phase entry cyclin 5.
GN CLB5 OR YPR120C OR P9642.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BF264-15D;
RC MEDLINE=92387544; PubMed=1387626;
RX Epstein C.B., Cross F.R.;
RA "CLB5: a novel B cyclin from budding yeast with a role in S phase.";
RL Genes Dev. 6:1695-1706(1992).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=AB320;
RC MEDLINE=93307652; PubMed=8319908;
RX Schwob E., Nasmyth K.;
RA "CLB5 and CLB6, a new pair of B cyclins involved in DNA replication
RT in Saccharomyces cerevisiae.";
RL Genes Dev. 7:1160-1175(1993).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=94074543; PubMed=8253070;
RX Kuehne C., Linder P.;
RA "A new pair of B-type cyclins from Saccharomyces cerevisiae that
RT function early in the cell cycle.";
RL EMBO J. 12:3437-3447(1993).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C / AB972;
RC MEDLINE=97313271; PubMed=9169875;
RX Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
RA Araujo R., Aparicio A., Barrall B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Dueterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Heblung U., Heumann K., Hilbert H., Hillier L.,
RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Klein K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
RA Scherrens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
RL Nature 387:103-105(1997).
CC -!- FUNCTION: REQUIRED FOR EFFICIENT PROGRESSION THROUGH S PHASE AND
CC POSSIBLY FOR THE NORMAL PROGRESSION THROUGH MEIOSIS. INTERACTS
CC WITH CDC28.
CC -!- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED JUST BEFORE CELL CYCLE
CC START.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M91209; AAA34503.1; -
CC EMBL; X70435; CA949893.1; -
DR EMBL; M91209; AAA34503.1; -
DR EMBL; X70435; CA949893.1; -

```

```

DR EMBL; U40828; AAB68061.1; -.
DR PIR; S31290; S31290.
DR SGD; S0006324; CLB5.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0000082; P:G1/S transition of mitotic cell cycle; IEP.
DR GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IEP.
DR GO; GO:0006279; P:premeiotic DNA synthesis; IGI.
DR InterPro; IPR006670; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR InterPro; IPR006671; Cyclin_N.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM00385; CYCLIN; 2.
DR PROSITE; PS00292; CYCLINS; 1.
DR CYCLIN; Cell cycle; Cell division; Multigene family.
KW CYCLIN; Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 435 AA; 50431 MW; 5AD67EB841BA5759 CRC64;
Query Match 24.4%; Score 53; DB 1; Length 435;
Best Local Similarity 46.4%; Pred. No. 7;
Matches 13; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
QY 3 SNLRLLNKAALRKAKHTSMVNFYKPK 30
DB 28 SNLKILQNKALSKNDSSSKQVQDSKP 55
RESULT 12
EFF1 CHLMU STANDARD; PRT; 185 AA.
AC Q9PKR6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor P1 (EF-P 1).
GN EFPI OR TC0398.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=MoPn / Nigg;
RC MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RL pneumoniae AR39.";
CC Nucleic Acids Res. 28:1397-1406(2000).
CC -!- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATES EFFICIENT
CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED
CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING
CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING
CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE
CC (BY SIMILARITY).
CC -!- PATHWAY: Protein biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the elongation factor P family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002306; AAF39255.1; -.
DR PIR; B81708; B81708.
DR TIGR; TC0398; -.
DR HAMAP; MF 00141; -.
DR InterPro; IPR001059; EF-P.

```

```

DR Pfam; PF01132; EFP; 1.
DR PROSITE; PS01275; EFP; 1.
KW Protein biosynthesis; Elongation factor; Complete proteome.
SQ SEQUENCE 185 AA; 20468 MW; 969BD834F4997CE2 CRC64;

Query Match 24.2%; Score 52.5; DB 1; Length 185;
Best Local Similarity 41.2%; Pred. No. 3.2;
Matches 14; Conservative 9; Mismatches 8; Indels 2; Gaps 2;

QY 11 KAALRKAKTSMV--RNFRYKPKVQSQVLKPRDL 42
| | | | | | | | | | | | | | | | | | | |
DB 38 KVALQAGSDVVVERFRKAGQEVK-EAQPEPRL 70

RESULT 13
NARX ECOLI
ID NARX ECOLI STANDARD; PRT; 598 AA.
AC P10956;
DT 01-JUL-1989 (Rel. 11, Created)
DI 01-AUG-1992 (Rel. 23, Last sequence update)
DI 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrate/nitrite sensor protein narX (EC 2.7.3.-).
GN NARX OR NARX OR B1222 OR Z1998 OR ECS1727 OR SF1225.
OS Escherichia coli.
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 623;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=89263708; PubMed=2657652;
RA Nohno T., Noji S., Taniguchi S., Saito T.;
RT "The narX and narL genes encoding the nitrate-sensing regulators of
RT Escherichia coli are homologous to a family of prokaryotic two-
RT component regulatory genes.";
RL Nucleic Acids Res. 17:2947-2957 (1989).
RN [2]
SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [3]
SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155 (1996).
RN [4]
SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grothbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).

[5] SEQUENCE FROM N.A.
RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11259796;
RA Hayashi T., Makino K., Ohnishi M., Kurakawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kunara S., Shiba T., Hattori M., Shinagawa H., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
RN [6]
SEQUENCE OF 1-55 FROM N.A.
RP SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=89338707; PubMed=2668029;
RA Noji S., Nohno T., Saito T., Taniguchi S.;
RT "The narX gene product participates in nitrate transport induced in
RT Escherichia coli nitrate-respiring cells.";
RL FEBS Lett. 252:139-143 (1989).
RN [7]
SEQUENCE OF 56-598 FROM N.A.
RP SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=89197802; PubMed=2649492;
RA Stewart V., Parales J. Jr., Merkel S.M.;
RT "Structure of genes narL and narX of the nar (nitrate reductase)
RT locus in Escherichia coli K-12.";
RL J. Bacteriol. 171:2229-2234 (1989).
RN [8]
SEQUENCE OF 1-60 FROM N.A.
RP SPECIES=E.coli; STRAIN=K12 / MC4100;
RX Cavicchioli R., Gunsalus R.P., Chiang R.C.;
RL Submitted (AUG-1992) to the EMBL/Genbank/DBSJ databases.
RN [9]
SEQUENCE FROM N.A.
RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang Q., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441 (2002).
CC -!- FUNCTION: ACTS AS A SENSOR FOR NITRATE/NITRITE AND TRANSDUCES
CC SIGNAL OF NITRATE AVAILABILITY TO THE NARL PROTEIN AND OF BOTH
CC NITRATE/NITRITE TO THE NARL PROTEIN. NARX PROBABLY ACTIVATES NARL
CC AND NARL BY PHOSPHORYLATION IN THE PRESENCE OF NITRATE. NARX ALSO
CC PLAYS A NEGATIVE ROLE IN CONTROLLING NARL ACTIVITY. PROBABLY
CC THROUGH DEPHOSPHORYLATION IN THE ABSENCE OF NITRATE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: Contains 1 HAMP domain.
CC -!- SIMILARITY: Contains 1 histidine kinase domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X13360; CAA31741.1; -
CC EMBL; AR000220; AAC74306.1; -
CC EMBL; D90757; BAA36090.1; -
CC EMBL; AE005339; AAG56082.1; -
CC EMBL; AF002556; BAB35150.1; -
CC EMBL; X15996; CAA34125.1; -
CC EMBL; M24910; AAA24198.1; -
CC EMBL; M65715; CAA46631.1; -
CC EMBL; X69189; CAA48934.1; -
CC EMBL; AS015149; AAN42838.1; -

```

```

DR PIR; F85702; F85702.
DR PIR; G30844; G30844.
DR PIR; S26137; RGEUNX.
DR EcoGene; EGI0646; narX.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR003661; His_kinA.
DR InterPro; IPR005467; His_kinase.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00388; HisK_A; 1.
DR PROSITE; PS50885; HAMP; 1.
DR PROSITE; PS50109; His_KIN; 1.
DR Sensory transduction; Transferrase; Kinase; Phosphorylation;
KW Transmembrane; Inner membrane; Nitrate assimilation;
KW Complete proteome.
FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 37 POTENTIAL.
FT DOMAIN 38 151 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 152 174 POTENTIAL.
FT DOMAIN 175 598 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 176 228 HAMP.
FT DOMAIN 393 597 HISTIDINE KINASE.
FT MOD_RES 399 399 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 42 52 GSAHLINKAGS -> AAPRSTKRDA (IN REF. 1 AND 6).
FT CONFLICT 374 374 A -> G (IN REF. 7).
SQ SEQUENCE 598 AA; 67083 MW; 886BA0FC2C8F3C3E CRC64;

Query Match 24.0%; Score 52; DB 1; Length 598;
Best Local Similarity 31.0%; Pred. No. 14;
Matches 13; Conservative 9; Mismatches 18; Indels 2; Gaps 1;

OY 1 MSLNRLILNKALRKAKHTSMVRNF--RYGKPVQSQVQLKPR 40
: : : : : : : : : : : : : : : : : : : :
Db 448 LITFRLQLTEGLRPALEASCEESYSAKEGFPVKLDYQLPFR 489

RESULT 14
POFB SCHPO STANDARD; PRT; 506 AA.
AC Q09855; Q9P7V1.
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE F-box/WD-repeat protein pof11.
GN POF11 OR SPAC29E6.01 OR SPAC30.05.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21668955; PubMed=11809834;
RA Katayama S., Kitamura K., Lehmann A., Nikaido O., Toda T.;
RT "Fission yeast F-box protein Pof3 is required for genome integrity and telomere function.";
RL Mol. Biol. Cell 13:211-224 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Bayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 81.5 kDa protein in HLU1-SMP2 intergenic region.
GN YMR163C OR YMR520.12C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Church C.M., Badcock K., Brown D., Chillingworth T.,
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
RT XIII."
RL Nature 387:90-93 (1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z49705; CAA89799.1; -.
DR PIR; S54521; S54521.
DR SGD; S0004773; YMR163C.
KW Hypothetical protein.
SQ SEQUENCE 705 AA; 81466 MW; E4F2A4D205A98F66 CRC64;
Query Match 23.5%; Score 51; DB 1; Length 705;
Best Local Similarity 36.8%; Pred. No. 23;
Matches 14; Conservative 5; Mismatches 19; Indels 0; Gaps 0;
QY 5 LRLLNKAALRKAAHTSMVRNFRYKPVQSQVQLKPRDL 42
DB 643 LRTAKSFELLRKQAQSMVXFGFKPLRDLDAFLSRPL 680

```

Search completed: December 3, 2003, 14:32:56
 Job time : 12 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:29:02 ; Search time 34 Seconds
(without alignments)
326.361 Million cell updates/sec

Title: US-08-765-244-22
Perfect score: 217
Sequence: 1 MSLNRLILNKALRAKHAHTS.....NFRYKPVQSQVQLKPRDL 43

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23:*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	199	91.7	102	11 P70518	P70518 rattus norv
2	164	75.6	351	11 Q8RIA8	Q8RIA8 mus muscullu
3	130	59.9	354	6 Q9N1U7	Q9N1U7 bos taurus
4	119	54.8	41	11 Q63786	KW Q63786 rattus norv
5	106	48.8	354	13 Q9IAU8	Q9IAU8 trachemys s
6	70	32.3	356	13 Q9IAV0	Q9IAV0 scelopor
7	61	28.1	1286	10 Q23351	Q23351 arabidopsis
8	60	27.6	351	17 Q9YH55	Q9YH55 sulfolobus
9	57	26.3	354	13 Q9YHY9	Q9YHY9 gallus gall
10	56	25.8	369	2 Q9Z635	Q9Z635 streptococ
11	55.5	25.6	454	10 Q9LWN3	Q9LWN3 arabidopsis
12	54	24.9	300	5 Q44679	Q44679 caenorhabd
13	54	24.9	457	13 Q8AVS9	Q8AVS9 xenopus lae
14	54	24.9	544	5 Q96570	Q96570 drosophila
15	54	24.9	866	5 Q9GPC3	Q9GPC3 drosophila
16	54	24.9	866	5 Q9W034	Q9W034 drosophila

17	54	24.9	1053	5 Q8IRH0	Q8IRH0 drosophila
18	54	24.9	1075	5 Q8IRH1	Q8IRH1 drosophila
19	53.5	24.7	141	16 Q8DG64	Q8DG64 synecococc
20	53.5	24.7	425	2 Q45965	Q45965 coxiella bu
21	53.5	24.7	755	5 Q9V4J5	Q9V4J5 drosophila
22	53.5	24.7	781	5 Q8IGP7	Q8IGP7 drosophila
23	53	24.4	247	10 Q9LIG9	Q9LIG9 oryza sativ
24	53	24.4	593	16 Q9ZC64	Q9ZC64 yersinia pe
25	53	24.4	601	16 Q8DOB5	Q8DOB5 yersinia pe
26	53	24.4	852	2 Q8KJM4	Q8KJM4 rhizobium l
27	52.5	24.2	140	12 Q9A1A4	Q9A1A4 bovine vira
28	52	24.0	213	16 Q9HX37	Q9HX37 pseudomonas
29	52	24.0	335	5 Q76719	Q76719 caenorhabd
30	52	24.0	357	13 Q9IAU9	Q9IAU9 alligator m
31	52	24.0	598	16 Q8FH22	Q8FH22 escherichia
32	52	24.0	1009	5 Q8MQJ6	Q8MQJ6 drosophila
33	52	24.0	3004	5 Q24550	Q24550 drosophila
34	52	24.0	3004	5 Q9VYN8	Q9VYN8 drosophila
35	51.5	23.7	260	16 Q50746	Q50746 borrelia bu
36	51	23.5	226	16 Q8EF28	Q8EF28 shewanella
37	51	23.5	376	5 Q8SUE7	Q8SUE7 encephalito
38	51	23.5	499	2 Q8KNN7	Q8KNN7 clostridium
39	51	23.5	598	16 Q8ZP35	Q8ZP35 salmonella
40	51	23.5	598	16 Q8Z7G1	Q8Z7G1 salmonella
41	51	23.5	633	16 P74554	P74554 synecocyst
42	51	23.5	645	2 Q9S603	Q9S603 bacillus th
43	51	23.5	1779	5 Q18150	Q18150 caenorhabd
44	50.5	23.3	138	16 Q9JUG9	Q9JUG9 neisseria m
45	50.5	23.3	296	5 Q9VXZ7	Q9VXZ7 drosophila

ALIGNMENTS

RESULT 1

P70518 PRELIMINARY; PRT; 102 AA.
AC P70518
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ornithine carbamoyltransferase precursor (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85051832; PubMed=6548714;
RA McIntyre P., Graf L., Mercer J., Peterson G., Hudson P.J.,
RA Hoogenraad N.;
RT "A highly basic N-terminal extension of the mitochondrial matrix
enzyme ornithine transcarbamylase from rat liver.";
RL FEBS Lett. 177:41-46(1984).
DR EMBL; K03041; AAA41771.1; -.
DR HSSP; P00480; 10TH.
DR InterPro; IPR006130; Asp/Orn COTranf.
DR InterPro; IPR006132; OTCace_P.
DR Pfam; PF02729; OTCace N; 1.
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
DR SIGNAL; Transferase.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 >102 ORNITHINE CARBAMOYLTRANSFERASE.
FT NON TER 102 102
SQ SEQUENCE 102 AA; 11932 MW; B739D41F19BF720D CRC64;

Query Match 91.7%; Score 199; DB 11; Length 102;

Best Local Similarity 97.6%; Pred.No. 9e-22; 1; Indels 0; Gaps 0;
Matches 41; Conservative 0; Mismatches 1;

QY 1 MSLNRLILNKALRAKHAHTSMVNFYKPVQSQVQLKPRDL 42
|||||
Db 1 MSLNRLILNKALRAKHAHTSMVNFYKPVQSQVQLKGRDL 42

Matches 22; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 MSLNRLILNKALRKAHTSMVNFYRGKPVQSQVQLKPRDL 42
 DB 1 MLFNRLSLFSTRNVNKLKSHLVQIRYRHGPPSETPVOLKGRDL 42

RESULT 6

Q9TAVO Q9TAVO PRELIMINARY; PRT; 356 AA.

AC Q9TAVO; EU Arabidopsis sequencing project;
 DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
 DE Ornithine transcarbamylase.
 GN OTC.
 OS Sceloporus undulatus (Eastern fence lizard) (skink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;
 OC Sceloporus.
 OX NCBI_TaxID=8520;
 RN [1]

SEQUENCE FROM N.A.

RP Shimogiri T., Koyanagi K., Mannen H., Tsuji S.;
 RA "Ornithine transcarbamylase is expressed in urocoelotic animals.";
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF134844; AAF61408.1;
 DR HSPF; P00480; IOTH.
 DR InterPro; IPR006130; Asp/Orn_Cotransf.
 DR InterPro; IPR002292; Orn_carbtransf.
 DR InterPro; IPR006131; OTCace_O.
 DR InterPro; IPR006132; OTCace_P.
 DR Pfam; PF00185; OTCace_1.
 DR Pfam; PF02729; OTCace_N; 1.
 DR PRINTS; P00100; AOTCASE.
 DR TIGRFAMs; TIGR00658; orn_carb tr; 1.
 DR PROSITE; PS00097; CARBAMOYLTRANSFEROSE; 1.
 SQ SEQUENCE 356 AA; 40399 MW; 01D49CCB93E4DBD1 CRC64;

Query Match 32.3%; Score 70; DB 13; Length 356;

Best Local Similarity 40.9%; Pred. No. 0.063;
 Matches 18; Conservative 6; Mismatches 18; Indels 2; Gaps 1;

QY 1 MSLNRLILNKALRKAHTSMVNFY--GKPVQSQVQLKPRDL 42
 DB 1 MLFNRLSLFSTRNVNKLKSHLVQIRYRHGPPSETPVOLKGRDL 44

RESULT 7

O23351 Q23351 PRELIMINARY; PRT; 1286 AA.

AC Q23351; EU Arabidopsis thaliana (Mouse-ear cress).
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE Hypothetical 143.9 kDa protein.
 GN AT4G14970.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euphorbia II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=98121113; PubMed=9461215;
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
 RA Bergkamp R., Dirks W., van Staveren M., Stiekema W., Drost L.,
 RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,
 RA Wedler E., Wambutt R., Weitzenecker T., Pohl T.M., Terry N.,
 RA Gleason J., Villarroel R., De Clerck R., van Montagu M., Lecharny A.,
 RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotter P.,
 RA Etian K.D., Rieger M., Schaeffer M., Funk B., Mueller-Auer S.,
 RA Silvey M., James R., Montfort A., Pons A., Puigdomenech P., Douka A.,

RA Vouklatou E., Milioni D., Hatzopoulos P., Piravandi E., Obermaier B.,
 RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,
 RA Falme K., Benes V., Rechman S., Ansgore W., Cooke R., Berger C.,
 RA Delsen M., Voet M., Voickaert G., Mewes H.W., Klosterman S.,
 RA Schueller C., Chalwatzis N.,
 RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
 Arabidopsis thaliana";
 RL Nature 391:485-488(1998).
 RN [2]

SEQUENCE FROM N.A.

RP EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z97337; CAB10276.1; --
 DR EMBL; AL161540; CAB78539.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 1286 AA; 143860 MW; F953E283C53D0DEE CRC64;

Query Match 28.1%; Score 61; DB 10; Length 1286;

Best Local Similarity 38.2%; Pred. No. 5.8;
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 3 SNLRLLNKALRKAHTSMVNFYRGKPVQSQV 36
 DB 107 SNLRRLSSSTTKRDESLVNLVSPQLDIQ 140

RESULT 8

Q97W55 Q97W55 PRELIMINARY; PRT; 351 AA.

AC Q97W55; EU Arabidopsis sequencing project;
 DT 01-OCT-2001 (TEMBLrel. 18, Created)
 DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Hypothetical GTP binding protein SSO2385.
 GN SSO2385.

OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]

SEQUENCE FROM N.A.

RP STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=1142726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
 RA Haikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Senses C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; A8006839; AAK42533.1; --
 DR InterPro; IPR004095; TGS_dom.
 DR Pfam; PF02824; TGS; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 351 AA; 39916 MW; 046A96BF004865DE CRC64;

Query Match 27.6%; Score 60; DB 17; Length 351;

Best Local Similarity 34.6%; Pred. No. 1.9;
 Matches 18; Conservative 8; Mismatches 10; Indels 16; Gaps 3;

QY 7 ILLNKAA-----LRKAHTSMVNFY----GKPVQSQV-----LKPRLD 42
 DB 294 LILKGGSTVLDVARKLHSSLAENFRYRVWVGSKVFGQGVGPHILEDRI 345

RESULT 9

Q9YH19 Q9YH19 PRELIMINARY; PRT; 354 AA.

ID Q9YH19; EU Arabidopsis sequencing project;
 AC Q9YH19;
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)

01-MAR-2003 (TReMBLrel. 23, Last annotation update)	
Ornithine transcarbamylase precursor (BC 2.1.3.3).	
OTC.	
Gallus gallus (Chicken).	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
Gallus.	
NCBI_TaxID=9031;	
[1]	
SEQUENCE FROM N.A.	
STRAIN=BRED WHITE LEGHORN; TISSUE=Kidney;	
MEDLINE=99011321; PubMed=9792920;	
RA Shimogiri T., Kono M., Mannen H., Mizutani M., Tsuji S.;	
RT "Chicken ornithine transcarbamylase gene, structure, regulation, and	
RT chromosomal assignment: repetitive sequence motif in intron 3	
RT regulates this enzyme activity.;"	
J. Biochem. 124:962-971(1998).	
DR EMBL; AF065629; AAD1234.1; -.	
DR EMBL; AF065638; AAD33083.1; -.	
DR EMBL; AF065630; AAD33083.1; JOINED.	
DR EMBL; AF065631; AAD33083.1; JOINED.	
DR EMBL; AF065632; AAD33083.1; JOINED.	
DR EMBL; AF065634; AAD33083.1; JOINED.	
DR EMBL; AF065635; AAD33083.1; JOINED.	
DR EMBL; AF065636; AAD33083.1; JOINED.	
DR EMBL; AF065637; AAD33083.1; JOINED.	
DR HSSP; P00480; IOTH.	
DR InterPro; IPR006130; Asp/Orn COTransf.	
DR InterPro; IPR002292; Orn.carbTransf.	
DR InterPro; IPR006131; OTCace O.	
DR InterPro; IPR006132; OTCace P.	
DR Pfam; PF00185; OTCace; 1.	
DR Pfam; PF02729; OTCace N; 1.	
DR PRINTS; PR00100; AOTCACE.	
DR TIGRfams; TIGR00658; ornithine carb. tr. 1.	
DR PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.	
DR transferase.	
SEQUENCE 354 AA; 40245 MW; 20447180BAD9D4D5D6 CR664;	

Query Match 26.3%; Score 57; DB 13; Length 354;
Best Local Similarity 40.5%; Pred. No. 5.5;
Matches 17; Conservative 4; Mismatches 21; Indels 0; Gaps 0;

Qy 1 MLNLRILLNKAALRKAHTSMVRNFYKGPVQSQVLKPRDL 42
||| : | | : | : | : | : | : |||
Dd 1 MLFNLNLYRITKLTONSKHLPRHFCRGEPNQMNVCVKGRDL 42

```

RESULT 10
Q92635 PRELIMINARY; PRRT; 369 AA.
ID Q92635
AC Q92635,
DC Q92635,
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative Glucosyl transferase.
OS CFS19CS.
OC Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_taxid=1313;
[1]
SEQUENCE FROM N.A.
RN RP MEDLINE=99395043; PubMed=10464207;
RX Morona J.K., Morona R., Paton J.C.;
RT "Comparative genetics of capsular polysaccharide biosynthesis in
RL Streptococcus pneumoniae types belonging to serogroup 19.";
RL J. Bacteriol. 181:5355-5364(1999).
DR EMBL; AF105116; AAD1925.1; -.
DR InterPro: IPR001296; Glyco trans.1.
DR Pfam: PF00534; Glycos_transf_1; 1.
DR transferase.
SQ SEQUENCE 369 AA; 43229 MW; FCFCFD5C106AC8BE CRC64;
KW

```

```

Query Match      25.8%; Score 56; DB 2; Length 369;
Best Local Similarity 32.5%; Pred. No. 8.1;
Matches 13; Conservative 9; Mismatches 14; Indels

QY 4 NLRILLNKAALFKAHTSMVRNFRYKPVQSQVQLKPRDLCC 43
:- : : : : : : : : : : : : : : : : : : : :
db 48 NVHKVLVRLGIKKSDMSMT----YTKAENVHLSPEVCC 83

```

```

RESULT 11.
Q91MN3 PRELIMINARY; PRT; 454 AA.
ID Q91MN3
AC Q91MN3;
DC DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DE F16P4.13 protein.
GN F16P4.13.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae
OC eucoids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
RX [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC Liu S.X., Sakano H., Yu G., Etgu P., Lee J.M., Lenz C., Pham P.,
RA Toriumi M., Byun E., Chan A., Chin C., Choi E., Chung M.,
RA Goldsmith A., Gonzalez A., Liu A., Smith A., Vaysberg M., Altafi H.,
RA Brooks S., Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,
RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,
RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
RA Federspiel N.A., Theologis A.;
RT "The sequence of BAC F16P4 from Arabidopsis thaliana chromosome 1."
RRL Submitted (JUL-2000) to the ENBL/GenBank/DBDB databases.
RQ ENBL; AC036104; AAF81364.1; -.
SQ SEQUENCE 454 AA; 50928 MW; 70BC2ACC9DBEDCF CRC64;

```

Query Match 25.6%; Score 55.5; DB 10; Length 454;
Best Local Similarity 45.2%; Pred. No. 12;
Matches 14; Conservative 5; Mismatches 11; Indels . 1; Gaps 1;

QY	11	KAALRKAHT-SMVNFRYKPVQSQVQLKPR	40
		: : : :	
D6	422	KAAYSRTSTDSPFNRYRESQPMGSPVQAFPR	452

RESULT	2
ID	O44679
PRELIMINARY;	PRT; 300 AA.
AC	O44679;
DT	01-JUN-1998 (TrEMBLrel. 06, Created)
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hypothetical 33.8 kDa protein.
GN	C14G6.13.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdita; Rhabditoidea;
OC	Rhabditidae; Pelodierae; Caenorhabditis.
OX	[1] NCBI_TaxID=6239;
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RY	MEDLINE=99069613; PubMed=9851916;
RA	None;
RT	"Genome sequence of the nematode C. elegans: a platform for
RT	investigating biology. The C. elegans Sequencing Consortium.";
RL	Science 282:2012-2018(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Bristol N2;
RA	David M., Wohlmann P., Bauer C., Antoniou B.;

RT "The sequence of *C. elegans* cosmid C14C6.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;
RA Waterston R;

KA Watsonson K.;
RT "Direct Submission.";
RL Submitted (JUL-2001) to the EMBL/genBank/DBJ databases.
DR EMBL; AF039051; AAB34289.1; -.
DR WormPep; C1406.13; CE08168.
DR

Query Match 24.9%; Score 54; DB 5; Length 300;
Best Local Similarity 31.4%; Pred. No. 13;
Matches 11; Conservative 10; Mismatches 14; Indels

Qy 2 LSNRLILNKAALRKAHTSMVRNFYRGKPVQSQVQ 36
 | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 199 LDNIRELCGOSALOKCSETRDPNTGFYRPINSEIO 233

RESULT 13

Q8AVS9 ID Q8AVS9 PRELIMINARY; PRT; 457 AA.

AC	Q8AVS9;	
DT	01-MAR-2003	(TrEMBLrel. 23, Created)
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)
DE		Similar to hypothetical protein MGC1193.
DE		<i>Xenopus laevis</i> (African clawed frog).
OS		<i>Xenopus laevis</i> (African clawed frog).
OS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC		Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC		Xenopodinae; Xenopus.
OX		NCBI TaxID=8355;

RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (DEC-2002) to the ENBL/GenBank/DBJ databases

Query Match 24.9%; Score 54; DB 13; Length 457;
Best Local Similarity 46.4%; Pred. No. 20;
Matches 13; Conservative 3; Mismatches 12; Indels

QY 2 LSNLRILLNKAALRKAHTSMWRNFRYK 29
| : | | : | | | | : | | | |
DB 135 LWDLRKLNKACTHGTHTSWVKNI EYDK 162

RESULT 14

RESULTS 14.
096570
ID 096570
PRELIMINARY:
PRT: 544 AA.

AC	O96570;	
AD		
AE		
AF		
AG		
AH		
AI		
AJ		
AK		
AL		
AM		
AN		
AO		
AP		
AQ		
AR		
AS		
AT		
AV		
AW		
AX		
AY		
AZ		
BA		
BB		
BC		
BD		
BE		
BF		
BG		
BH		
BI		
BJ		
BK		
BL		
BM		
BN		
BO		
BP		
BQ		
BR		
BS		
BT		
BU		
BV		
BW		
BX		
BY		
BZ		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CA		
CB		
CC		
CD		
CE		

EN	[a] -	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BEIRUT;	
RX	MEDLINE=99250256; PubMed=10231575;	
RA	Tarenkov A., Saez A.G., Ayala F.J.;	
RT	"A compact gene cluster in <i>Drosophila</i> : the unrelated <i>Cs</i> gene is	

Query Match	24.9%;	Score 54;	DB 5;	Length 544;
Best Local Similarity	34.2%;	Pred. No. 25;		
Matches 13:	Conservative	5;	Mismatches	20;
				Indels

QY 1 MLNLRILLNKAAALRKAHTSMVRNFYKGPVQSQQVLK 38
:
: | | : | | | : | | : | | : |
db 332 VLKNFSAILFKPALPLEKLOAIRNLGYGNPKVIYLAYK 369

RESULT 15

RESOL 13
Q9GPG3
ID Q9GPG3 PRELIMINARY: PRT: 866 AA.

AC Q3GPG3; 2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Puromycin-sensitive aminopeptidase.
DE PSA OR CG1009
GN Drosophila melanogaster (Fruit fly).
OC Rosophila
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulz C., Peretzgaaga L., Fuller M.T.;
RT "Functional analysis of a *Drosophila* pyromycin-sensitive
RT aminopeptidase gene."; [PMID:11511111](#)
RL Submitted (DEC-2000) to the ENBL/GenBank/DBJ databases.

Query Match 24.9%; Score 54; DB 5; Length 866;
Best Local Similarity 33.3%; Pred. No. 42;
Matches 15: Conservative 11; Mismatches 7; Indels

QY 2 LSNRLILNKAAALRKAHTSMTVNF-RYGK-----PVQSQVLKPRD 41
| : | | | : | : | : | : | : | : | : | : | : | : |
613 TNNLHIT-----SHTRIMETHEPCEPNIVEDVAVPLGMPDPD 650

Search completed: December 3, 2003, 14:33:45
Job time : 36 secs